Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics

Simicevic, Jovan ; Schmid, Adrien W ; Gilardoni, Paola A ; Zoller, Benjamin ; Raghav, Sunil K ; Krier, Irina ; Gubelmann, Carinne ; Lisacek, Frédérique ; Naef, Felix ; Moniatte, Marc ; Deplancke, Bart

Abstract: The cellular abundance of transcription factors (TFs) is an important determinant of their regulatory activities. Deriving TF copy numbers is therefore crucial to understanding how these proteins control gene expression. We describe a sensitive selected reaction monitoring–based mass spectrometry assay that allowed us to determine the copy numbers of up to ten proteins simultaneously. We applied this approach to profile the absolute levels of key TFs, including PPAR and RXR , during terminal differentiation of mouse 3T3-L1 pre-adipocytes. Our analyses revealed that individual TF abundance differs dramatically (from 250 to >300,000 copies per nucleus) and that their dynamic range during differentiation can vary up to fivefold. We also formulated a DNnA binding model for PPAR based on TF copy number, binding energetics and local chromatin state. This model explains the increase in PPAR binding sites during the final differentiation stage that occurs despite a concurrent saturation in PPAR copy number.

DOI: https://doi.org/10.1038/nmeth.2441

Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics

Jovan Simicevic[1,5], Adrien W. Schmid[2,5], Paola A. Gilardoni[1,5], Benjamin Zoller[3], Sunil K. Raghav[1], Irina Krier[1], Carine Gubelmann[1], Frédérique Lisacek[4], Felix Naef[3], Marc Moniatte[2,*], Bart Deplancke[1,*]

[1]Laboratory of Systems Biology and Genetics, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[2] Proteomics Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[3] Laboratory of Computational Systems Biology, Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[4]SIB Swiss Institute of Bioinformatics, Geneva, Switzerland

[5]These authors contributed equally to this work

*Corresponding authors:

Marc Moniatte
EPFL Proteomics Core Facility
EPFL-PCF-PTP, Station 15
CH-1015 Lausanne
Tel: +41 (0)21 693 17 53
Fax: +41 (0)21 693 18 88
E–mail: marc.moniatte@epfl.ch

Bart Deplancke
Laboratory of Systems Biology and Genetics (LSBG)
EPFL-SV-IBI-LBSG, Station 19
CH-1015 Lausanne
Tel: +41 (0)21 693 18 21
Fax: +41 (0)21 693 09 80
Email: bart.deplancke@epfl.ch

Abstract

The regulatory properties of transcription factors (TFs) are largely dictated by their cellular abundance. Deriving absolute TF copy numbers is therefore crucial to understand how these proteins control gene expression. Here, we implement a sensitive selected reaction monitoring (SRM)-based mass spectrometry assay, allowing us to simultaneously determine the copy numbers of up to 10 proteins. We apply this approach to profile the levels of key TFs during terminal differentiation of mouse 3T3-L1 pre-adipocytes, revealing that TF abundance differs dramatically (from ~250 to >300,000 copies per nucleus), but that their dynamic range during differentiation varies at most five-fold. We also formulate a genome-wide TF DNA binding model to explain the significant increase in binding sites of the adipogenic master regulator PPARγ during the final differentiation stage, despite a concurrent saturation in PPARγ copy number. This model provides unique, quantitative insights into the relative contributions of binding energetics, copy number, and chromatin state in dictating TF DNA occupancy profiles.

Introduction

Differential gene expression is controlled by gene regulatory networks, which consist of functional interactions between regulatory state (e.g. transcription factors (TFs) and genomic (e.g. gene promoters, enhancers) components[1]. A major current interest is to derive models of gene regulatory networks that elucidate or predict the dynamic transcriptional mechanisms underlying complex gene expression programs[2-4]. An important drawback of existing models is that relatively little quantitative data has so far been used in their calibration. In particular, accurate measurements on the absolute, molecular abundance of most TFs are still very sparse[2], even though such information is key to understand most biochemical and regulatory processes involving this type of proteins[5, 6]. This conundrum can be explained by the relatively low expression of TFs in cells[7], which make their direct quantification or that of any other low abundant protein by mass spectrometry (MS) or other assays a substantial challenge.

Newly emerging proteomic approaches combining selected reaction monitoring (SRM) or high resolution MS with isotope dilution quantification strategies now promise to alleviate this scarcity in copy number data as they enable quantitative analyses over a wide concentration range[8-10]. These approaches all share the same principle of targeting a subset of detectable peptides, which are specific to the protein of interest (proteotypic peptides)[11]. Quantification is then achieved by comparing their extracted MS signals to those of accurately quantified, isotopically-labeled peptides, which are used as internal standards[8]. These peptides can be produced synthetically[8] or are now increasingly being derived from the proteolysis of full-length proteins that are themselves recombinantly[12-16] expressed or synthetically derived[17]. This is because the use of isotopically labeled full-length proteins offers clear advantages over other recently developed quantification approaches (a comparative overview of these approaches is shown in **Supplementary Table 1**). Both synthetic- and recombinant-based approaches offer the possibility of introducing single reference (e.g. [16-19]) or solubility (e.g. albumin binding protein[15] or glutathione-S-transferase[14]) tags linked to parent proteins. This enables the stoichiometric quantification of peptide standards released upon protein digestion, which are then used to quantify the endogenous protein within a complex sample. However, despite their potential in deriving accurate protein copy number measurements, particularly for low abundant proteins, only a few studies incorporating such targeted MS

approaches have so far demonstrated their capacity to accurately measure copy numbers of TFs in higher eukaryotes[14, 15]. In addition, none of these studies were to our knowledge specifically optimized to tackle the difficult task of monitoring the dynamic expression of TFs over the course of specific biological processes. Furthermore, currently available approaches typically rely on the use of a single reference peptide or tag which is spiked early in the analytical process. While the latter corrects for the concomitant loss in endogenous peptides during sample digestion and fractionation and thus increases accuracy and reproducibility[20, 21], the use of only one reference tag requires that the quantification of each protein standard is accomplished separately. This ultimately adds avoidable steps into the workflow potentially introducing undesirable artifacts (**Supplementary Table 1**).

These collective considerations prompted us to develop an SRM-based workflow, specifically tailored around the quantification of more lowly abundant proteins such as TFs, in which the internal protein standard quantification step is entirely performed *in situ* (i.e. the protein standard and its endogenous counterpart are simultaneously quantified). Here, we applied this workflow to assess the dynamic changes in absolute TF levels during the terminal phase of adipogenesis. While quantitative proteomic studies have been performed to explore the adipogenic proteome, these lacked the required sensitivity to detect and quantify even core adipogenic master regulators such as PPARγ and RXRα[22, 23], illustrating the inherent complexity of targeting TFs[23].

Results

Selection of TF-specific transitions for targeted SRM

The selection of proteotypic peptides for TFs proved challenging, mainly due to data scarcity in public repositories and to difficulties in detecting tryptic peptides in discovery experiments (**Supplementary Fig. 1** and **Online Methods**). We therefore decided to adopt an *in vitro* full-length protein expression-based strategy to derive proteotypic peptides, which allows for a better accounting of the local environment surrounding the peptide cleavage sites (**Online Methods** and data not shown). Ten TFs (including PPARγ and RXRα) that are expressed during terminal adipogenesis were selected after which the corresponding recombinant proteins were tagged with C-terminal glutathione S-transferase (GST) for full-length expression validation and purification (**Supplementary Table 2** and **Supplementary Fig. 2)**. In-gel separation

followed by high resolution LC-MS/MS on the tagged TF-containing fraction then enabled us to generate a collection of experimentally detected TF-specific tryptic peptides for the 10 proteins (**Supplementary Data 1** and **Online Methods**).

Implementation of the SRM-based methodology

Given the benefits of using full-length proteins as standards over other sources[12, 24] (**Supplementary Table 1**), we set out to utilize isotopically-labeled, *in vitro*-expressed TFs directly as standards for quantification of their endogenous counterparts. Building on pioneering work[18, 19, 25], we implemented a modified version of the same workflow that we have adopted for transition selection. For this purpose, we aimed to add an in-frame proteotypic reference-peptide tag at the N–terminus of the protein constructs that should serve as surrogate during the quantification step. After having considered several possibilities[18, 19] (**Online Methods**), we found that the SH-Quant peptide[19] offers optimal solubility and stability properties (**Supplementary Fig. 3**). We therefore adopted it for the *in situ* quantification of heavy TFs utilizing a synthetic and accurately quantified light SH-Quant counterpart. In this manner, one single peptide only (the SH-Quant) is used for the absolute quantification of a heavy, *in vitro*-expressed standard TF, thus simplifying and reducing the cost of the assays even further. In turn, quantification of the endogenous TF in absolute amounts is achieved by utilizing TF-specific peptides (**Fig. 1a-b**).

Absolute quantification of PPARγ and RXRα

To evaluate PPARγ and RXRα peptide transitions, two separate pilot SRM assays were implemented as depicted in **Fig. 1a** and exemplified in **Fig. 1b-c** (details in **Supplementary Figs. 4-6**). Subsequently, the complete workflow was applied to monitor the absolute, nuclear abundance of PPARγ and RXRα during terminal adipogenesis. The technical robustness of the whole workflow is highlighted by the low coefficient of variation (CV) obtained within one biological sample (**Supplementary Fig. 7**). The resulting data, as summarized in **Fig. 2** and **Supplementary Table 3**, provide unique insights into the dynamic protein copy number per cell (nucleus) profiles of PPARγ and RXRα.

To corroborate our results and to provide an independent assessment of the accuracy of our method, we performed three validation experiments. First, we performed Western blot analysis for each time-point (**Supplementary Fig. 8** and **Online Methods**), revealing that the detected TF abundance profiles mirror those observed by our SRM-based quantification approach. To provide a more quantitative assessment of the validity of our data, we performed semi-quantitative Western blotting on RXRα. Given the unavailability of purified mouse RXRα, we generated a calibration curve utilizing SRM for the quantification of *in vitro*-expressed RXRα standard and used it for Western blot-based quantification of endogenous RXRα at different time points during terminal adipogenesis (see **Supplementary Fig. 9a**-**b** and **Online Methods**). The resulting data show a good correlation with the SRM-based measurements (**Supplementary Fig. 9c**-**d,** R^2 = 0.72), although we found that the Western analysis tends to inflate RXRα copy numbers as compared to SRM-derived data.

To resolve this finding and due to the not truly independent nature of this experiment, we performed a second validation assay. For this purpose, we randomly selected and accurately quantified using the same workflow presented in **Fig. 1a** two human proteins of different molecular weight, α-synuclein (SYN, 14.5 kDa) and histidyl-tRNA synthetase (HARS, 57.4 kDa), to assess overall protein recovery. The HARS protein was selected as it falls within the molecular weight range of our TFs of interest, whereas SYN was selected to test the recovery of a significantly smaller protein. Each protein was spiked in amounts corresponding to 5,000 and 50,000 copies per cell into nuclear extracts to mimic the concentration range observed for TFs as well as the experimental nucleic background proteome (**Online Methods**). Our SRM validation predicted 4,620 and 45,950 copies of HARS per cell corresponding to an overall protein recovery of ~92% for both (**Supplementary Figs. 10a-e** and **11**). Slightly lower numbers were obtained for SYN, showing a recovery ranging from 65% to 80% (**Supplementary Figs. 10f** and **11**). This lower performance may be caused by several factors, although we suspect that its small size may have led to greater diffusion during the gel de- and re-hydration processes applied prior to digestion, resulting in greater protein loss.

Finally, to assess the impact of gel fractionation on the accuracy of our method, we performed a relatively straight-forward gel recovery experiment using

RXRα as reference protein (see **Online Methods**). As can be observed in **Supplementary Fig. 12**, we found that losses associated with sample processing were in the range of 30% for small protein amounts (<2.5 ng), whereas this improved to only ~8% when dealing with larger amounts. It is reasonable to conclude that absolute values obtained for the targeted TFs at different differentiation times should lie within the same error margin.

Together, these experiments provide a candid assessment of the precision of the presented SRM-based quantification approach, revealing a high accuracy for protein quantities above 2.5 ng based on the HARS and RXRα measurements. These values are in line with those reported in recent studies (e.g.[15]) and are thus consistent with what is considered an acceptable measurement inexactitude in the field. In addition, these experiments demonstrate that our pipeline is in principle applicable to any protein as long as it expresses well *in vitro*.

A quantitative model of genome-wide TF DNA binding

The availability of absolute protein copy number data allowed us to examine the relationship between the number of genome-wide DNA binding events and TF molecules. Using published RXRα and PPARγ ChIP-seq data[26], we found that, until the fourth differentiation day, the number of TF molecules and corresponding binding events correlate well (R^2 = 0.96 for PPARγ; R^2 = 0.85 for RXRα; **Supplementary Fig. 13**), and that in general, there are substantially more RXRα than PPARγ binding events consistent with their corresponding TF copy numbers. These high correlations indicate that ChIP-seq occupancy data likely reflect endogenous conditions rather than being the result of differences in antibody-mediated protein recoveries as previously hypothesized[26]. A striking discrepancy was however observed for day six as, compared to previous days, the number of RXRα and PPARγ binding events significantly increases (>three-fold for both TFs), whereas the number of TF molecules saturates or even decreases (**Supplementary Fig. 13**).

To reconcile these findings and to possibly provide a first-principle and quantitative analysis of this phenomenon, we generated a quantitative model to predict the number of PPARγ binding events in the genome given a number of TF molecules, ChIP detection threshold, and genome accessibility maps (see **Supplementary Note** for more details). This model bears similarity to previous

thermodynamic models of protein-DNA interactions[3, 27, 28], but also features some important differences, notably through the use of absolute TF levels instead of inferred protein concentration data and the focus on modeling DNA binding behavior over the course of a mammalian cellular differentiation process. To incorporate genome accessibility data into our model, we used published genome-wide data of the chromatin mark histone three lysine 27 mono-acetylation (H3K27Ac) for differentiation days zero, two, and six[29], since this mark is enriched at active and thus accessible genomic regions. Consequently, we formulated the model to account for the distribution of specific and non-specific sites for PPARγ in H3K27Ac-enriched regions (**Fig. 3a**). To assess the DNA binding events at thermodynamic equilibrium, we assumed that all the PPARγ proteins were on the DNA at either specific or non-specific sites, consistent with the consensus in the field that TFs tend to mostly reside on DNA *in vivo*[6]. For the sake of simplicity, we did not consider interactions with other TFs, or hindrance due to other proteins sitting on DNA (for a discussion on these topics, see **Supplementary Note**).

The resulting model predicts the number of binding events given the number of PPARγ copies per cell (nucleus) for each time point as a function of the ChIP detection threshold expressed as percentage occupancy (or residency time) (**Fig. 3b**). While it is difficult to precisely estimate the latter threshold, a survey of our own and published ChIP data revealed that the signal associated with DNA binding is typically considered as positive from around 1% (as compared to input)[30, 31]. Interestingly, when we used ~1% as detection threshold, our model predicts binding event numbers which closely mimic the experimentally derived data (**Fig. 3c**). Thus, the model correctly predicts the temporal pattern in the number of binding sites over the course of terminal adipogenesis.

As an independent test, we also repeated our analyses using another PPARγ ChIP-seq data set[32], which yielded similar conclusions (**Supplementary Note**). A closer look into the determining factors revealed that the sharp increase in binding events at day six as compared to days zero and two is in part driven by the increase in PPARγ copy number between days two and six. However, we found that an equally important factor relates to shifts in the relative distributions of higher versus lower affinity sites in accessible genomic regions at the different differentiation time points (**Fig. 3a**). Specifically, whereas the overall size of accessible regions is similar

across all days, we found that cells undergo important chromatin remodeling between days six as compared to zero and two such that more medium-to-high affinity sites for PPARγ become available. This in turn allows for a substantial increase in detectable DNA binding events even though the TF copy number increase is relatively modest. Together, these analyses show the value of accurately quantifying TF copy numbers to generate quantitative DNA binding models from which emergent properties regarding gene regulatory mechanisms underlying a specific biological process can be derived.

Multiplexing the SRM assay

Next, we sought to overcome the current limitation of quantifying only one protein per assay. For this purpose, we designed nine quantotypic SH-Quant tag variants that, along with the original SH-Quant tag, allow the quantification of up to 10 proteins or TFs in one single integrated SRM assay. We subsequently generated 10 expression vectors enabling the coupling of each tag variant to a distinct, adipogenic TF (among which we also included PPARγ and RXRα as validation). Corresponding transitions as well as the retention time of the parent ions are presented in **Fig. 4a**. The performance as well as chromatographic separation of the nine, newly designed candidates and their parent SH-Quant sequence was tested in a series of SRM runs (**Fig. 4a-e**). Two SH-Quant tag-variants "AADITSLYK" and "AAEVTSLYK" are isobaric and yet could be clearly distinguished by their chromatographic profile and by their distinct transitions (**Fig. 4b-e**)**.** A standard curve was established for each SH-Quant tag variant by spiking them into the same 3T3-L1 nuclear extract preparation that was next used for TF abundance measurements (**Supplementary Figs. 14** and **15**). This allowed us to determine the lower limit of quantification (LLOQ) for each tag (**Supplementary Fig. 16**). Applying the multiplex version of our methodology, we were able to monitor endogenous levels of 10 TFs (**Supplementary Table 2**) in 3T3-L1 nuclear extracts derived from three terminal differentiation time points (Day 0, Day 2, and Day 4) (**Fig. 5a-b**). As summarized in **Fig. 5c** and **Supplementary Fig. 17**, we found that the majority of the additionally analyzed TFs (e.g. PIAS4, NFIB, SMAD2, RARγ, FOSL2) are expressed within approximately the same copy number range as PPARγ or RXRα (**Supplementary Table 3**). Importantly, measurements of PPARγ and RXRα copy numbers were

coherent with data obtained with the non-multiplex approach, indicating that the up-scaling of the SRM assay does not interfere with measurement precision. The abundance of other TFs such as PIAS3, NR2C1, and ARID3a fell below the quantification threshold. Nevertheless, the overall quality of the fragment ions monitored for the heavy and light peptides (**Supplementary Fig. 18**) clearly allows for the unambiguous identification of these TFs.

Discussion

Determining the copy number of TFs has been of longstanding interest given that the DNA binding ability of a TF and hence its regulatory capacity strongly depends on its cellular concentration[5, 6]. However, despite the fundamental importance of TFs in most biological processes, a recent survey showed that the number of studies that have so far provided estimates on the abundance of animal TFs is low (~25)[6]. These estimates were moreover mostly achieved with indirect immuno-based measuring methods whose additional drawback is the dependence on antibodies which are available for only a low number of TFs. Nevertheless, these studies estimated TFs to be expressed in the range of 5,000 to several hundred thousand copies per nucleus[6] in line with our results.

To increase our understanding of TF behavior and to circumvent the issues typically encountered with immuno-based methods, we developed an SRM-based workflow which affords sufficient sensitivity and analytical reproducibility to generate quantitative data in absolute terms. This workflow offers several advantages:

First, the only requisite of our SRM-based assay is the relatively straight-forward *in vitro* expression of the protein candidate, which, except for certain types of proteins, becomes relatively trivial given the availability of efficient cloning and expression systems and Open-Reading Frame (ORF) clone libraries[33, 34]. Having easy access to the full-length protein has thereby the advantage of facilitating the empirical selection of the best quantotypic peptides for protein types that may not be well represented in spectral library databases (such as TFs). In addition, the ability to nearly extemporaneously produce heavy-labeled full-length protein standards, and to simultaneously quantify these protein standards and their endogenous counterparts within the same assay bypasses another significant source of variability in any quantitative assay, namely the necessity for protein/peptide storage[24].

Second, the assay is pillared on the use of the quantotypic SH-Quant reporter peptide which was found to feature high digestability, excellent solubility, ionisability and stability, and thus to exhibit optimal peptide properties for the robust and accurate quantification of proteins. Importantly, the SH-Quant tag allowed the design of tag variants with very similar properties enabling the *in situ* analysis of multiple proteins in one single assay in robust, sensitive, and reproducible fashion. So far, we have deliberately limited the number of TFs that are simultaneously quantified *in situ* for this assay. This is because we found that further multiplexing may jeopardize the overall sensitivity of the assay as the latter was specifically conceived to accommodate such throughput in the context of quantifying TFs. However, our approach is in principle not limited to 10 target proteins, providing that additional quantotypic tags are designed and validated and that the same overall accuracy and sensitivity compared to the "one protein at a time" approach is maintained. Together, we believe that our quantification strategy constitutes a powerful alternative to quantitative immunoassays (immunoblotting or ELISA) to determine fluctuations in protein abundance over a wide concentration range with the additional advantages over such assays of being able to multiplex and being independent of protein-specific antibodies[35]. This multiplex capacity should be of great value to those interested in examining or modeling absolute, dynamic changes of entire pathways or sets of biomarkers of interest in wildtype versus clinical, disease, or perturbation settings.

Here, we illustrate the importance of deriving this type of data by building a quantitative model aiming to clarify an intriguing discrepancy between TF binding site and copy number data. As such, we provide a quantitative explanation for the common observation that many TFs, including PPARγ, bind to significantly fewer sites in the genome than predicted based on the presence of their respective consensus motifs[29, 36, 37]. Whereas it was suggested that other factors may contribute to this binding site selectivity, here, we demonstrate that, at least for PPARγ, its DNA binding profile can be closely modeled by simply considering its own copy number, simple thermodynamic principles, and chromatin accessibility. The functional consequence of our findings is that the chromatin state constitutes a landing map for PPARγ DNA binding, consistent with the emerging notion of the importance of chromatin structure in shaping TF DNA binding patterns[37].

Acknowledgements

We would like to thank M. Lai (Thermo Fischer Scientific) for his help on solving Pinpoint software-related issues; M. Gstaiger (ETHZ) and H. Steen (Boston Children's Hospital) for providing us with the SH-Quant and the FLEX peptide tags respectively. H. Lashuel (EPFL) for providing us with the α-synuclein clone; E. Ahrné (University of Basel) for computational assistance and providing feedback on the manuscript together with S. Waszak (EPFL), M. Mueller (SIB, University of Geneva), and A. Schmidt (University of Basel); D. Chiappe for technical and F. Armand (all EPFL) for computational assistance. This work has been supported by funds from the Swiss National Science Foundation, by a Marie Curie International Reintegration Grant (to B.D.) from the Seventh Research Framework Program, by a SystemsX.ch iPhD Fellowship (to J.S.) and Grant (CycliX), by the Roland Bailly Foundation (Geneva, CH), by the NCCR Frontiers in Genetics Program, by EMBO (ALTF 459-2012) and European Commission (EMBOCOFUND2010, GA-2010-267146) support from Marie Curie Actions (to P.G.), and by Institutional support from the Ecole Polytechnique Fédérale de Lausanne (EPFL).

Author Contributions

Conceived and planned the study: J.S., A.S., P.G., M.M., B.D. Prepared the manuscript: J.S., A.S., P.G., B.Z., F.N., M.M., B.D. Performed wet bench experiments: J.S., A.S., P.G., S.R., C.G. Mass spectrometry data analysis: A.S. Biological data analysis: J.S., P.G., B.D. Modeling: B.Z., I.K., F.N., B.D. All the authors discussed the results and commented on the paper.

Figures

Figure 1: Workflow for the absolute quantification of TFs in 3T3-L1 cells.

(**a**) The left panel shows the preparation of 3T3-L1 total nuclear protein extract (NE). Cells are lysed at the indicated differentiation time point (Day0 (D0) to Day 6 (D6)) after which nuclear proteins are extracted. The resulting protein mixture is separated by SDS-PAGE and bands at the migration height of TF's are excised from the gel. The far right column depicts the preparation of *in vitro*-expressed SH-tagged TFs. The constructs are expressed in their heavy-labeled version (*) using a wheat germ-based *in vitro* transcription-translation kit, purified by GST affinity and separated by SDS-PAGE. Bands containing the heavy-labeled constructs, here SH-RXRα-GST*, are excised from the gel. Each nuclear extract band to be quantified is mixed with a gel slice of the *in vitro*-expressed TF construct, spiked with known amounts of light SH-Quant tag (center), and in-gel digested together. The resulting peptide mixtures are then quantified by SRM using validated proteotypic peptides selected from information previously collected on each *in vitro*-expressed TF via shotgun-MS runs. Each TF quantification requires a separate experiment in this configuration. (**b**) A schematic presentation of the quantitation approach as outlined in (**a**). (**c**) i. RXRα protein sequence with its highlighted tryptic peptides sequences used for quantitation. ii. SRM chromatogram of the five RXRα proteotypic fragments and its SH-Quant tag. iii. Zoom of the SH-Quant peptide in its light and heavy isoform and the Pinpoint software calculated curves for both isoforms (iv).

Figure 2. Summary of RXRα and PPARγ levels quantified by SRM.

(**a**) A graphical summary of the RXRα (left) and PPARγ (right) concentrations, expressed as fmol μg^{-1} nuclear extract (NE), in three individual biological replicates. (**b**) Absolute RXRα and PPARγ levels visualized as copies per cell. The bars represent the mean ± s.e.m of three biological replicates (*n*=3).

Figure 3. Quantitative modeling of genome-wide PPARγ DNA binding.

(**a**) Cumulative number of genomic sites in H3K27Ac regions at Day 0 (gray line), Day 2 (dashed line), and Day 6 (black line). The axis runs from the strongest, i.e. 1000-fold stronger than non-specific sites, to medium affinity sites, i.e. 60-fold

stronger. While the numbers for Day 0 and Day 2 are comparable, Day 6 shows a 35% increase in this range. (**b**) Number of detected bound loci at Day 0 (gray line), Day 2 (dashed line), and Day 6 (black line) during 3T3-L1 terminal differentiation in function of the detection threshold on the expected occupancy. The model takes into account the measured PPARγ copies per cell (nucleus) and the distribution of accessible high affinity sites. Matching with the number of experimentally derived PPARγ ChIP-seq sites predicts an occupancy threshold of 1.35%. (**c**) Temporal pattern for the predicted number of detected PPARγ-bound sites (dashed line), the actual number of measured sites (black line), and the protein copy number (grey line). Note that the number of sites shows an exponential-like increase, while the protein copy number graph reflects saturation.

Figure 4. Properties of the SH-Quant tag variants.

(**a**) The table provides a detailed summary of the different peptide-tags specifically designed for the multiplex assay. The amino acid sequences of tags 2 through 10 were derived from the original SH-Quant tag (1), and were designed with the intent of conserving intense y" fragment ions at positions y2 to y5 for quantification purposes. The fragmentation fingerprint of modified ions y6 to y8 allows for improved tag identification, particularly when dealing with isobaric peptides. This specific peptide design allows for a clear LC separation of the isobaric tags: 3 versus 7 (m/z = 470.2; green), and 1 versus 9 (m/z = 491.2; red). (**b**) LC-SRM chromatogram showing the extent of separation of the peaks belonging to the SH-Quant peptide-variants; proper separation was achieved for most tags using a short LC gradient of 60 min. (**c**) LC-SRM chromatogram of a multiplexed nuclear extract sample using a 120 minute gradient. (**d**) Zoom-in view of the area highlighted in panel c), showing the extracted ion current (EIC) of isobaric SH-Quant tags 1 and 9 in a multiplex sample (peak 1: m/z = 491.2, peak 2: m/z = 491.2). (**e**) Calculation of the peak areas of the light and heavy peptide counterparts using Pinpoint. The difference in physicochemical properties of two isobaric SH-Quant tags 9 and 1 result in separate elution times, 30.7 min (AAEVTSLYK) and 35 min (AADITSLYK) respectively. Identical elution times were observed for both light and heavy counterparts in all 10 SH-Quant tags.

Figure 5. Simultaneous monitoring of the nuclear abundance of ten TFs during terminal adipogenesis via spiking of *in vitro*-expressed full-length TFs.

(**a**) Multiplex SRM transition profiles of selected best-responding tryptic peptides from all ten TFs (light and heavy) during a 120 minute LC-SRM run. (**b**) Zoom-in (minutes 29-36) of the highlighted region in spectrum (**a**), illustrating the complexity of the mixture and the separation quality achieved with the multiplex SRM assay on nuclear extract sample. (**c**) Calculated endogenous levels of all ten TFs detected at Day 0, Day 2, and Day 4 (bars represent the mean ± s.e.m of 4-6 technical replicates (*n*=4-6)).

Online Methods

3T3-L1 cell culture, differentiation, and protein extraction

3T3-L1 cells were cultured and differentiated into adipocytes as detailed in Raghav *et al.,*[31]. Briefly, 3T3-L1 cells were collected at six different differentiation time points (0h, 2h, Day 1, Day 2, Day 4, and Day 6). At each time-point, petri-plates were rinsed twice with 1X PBS, after which cells were trypsinized, rinsed once with cold 1X PBS, and centrifuged and the pellets were then stored at –80°C. Cells were lysed in cold lysis buffer (10mM HEPES-NaOH at pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT, 0.5% NP-40, 0.1 mM EDTA, 0.1 mM EGTA) containing protease inhibitors and phosphates inhibitors (Roche) for 15 min and centrifuged for 10 min at 6,000 rpm (at 4°C) to sediment the nuclei. The pellet was washed twice to remove non-nuclear particles. The isolated nuclei were then washed using protein extraction buffer (20 mM HEPES-NaOH at pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.1 mM EDTA, 0.1 mM EGTA) after adding protease inhibitor and phosphatase inhibitor cocktail tablets (Roche) at 4°C for 30 min and centrifuged for 10 min at 6,000 rpm (at 4°C). Protein concentration was measured for each time point utilizing the Quick Start Bradford Dye Reagent (Bio-Rad). The supernatant containing nuclear proteins was collected and stored at –80°C.

Cloning and plasmids

Mouse TF ORFs were cloned in Gateway format essentially as described earlier[31, 34]. To make the wheat germ (WG) *in vitro* transcription-translation expression system (Promega) compatible with the mouse TF ORF clones and to allow the purification of TFs, we modified the SP6 pF3A WG (BYDV) Flexi vector (Promega) to accommodate the Gateway reading frame A cassette (Invitrogen). A glutathione-S-transferase (GST)-coding sequence containing a stop codon was subsequently incorporated in frame at the 3'-end of the second (i.e. 3') Gateway site using standard cloning techniques. TFs were then subcloned from the Gateway entry clone level into this "in-house" modified pF3A-GST destination vector by standard Gateway cloning.

In vitro protein expression and purification, gel staining, and immunoblotting

For selection of optimal TF transitions: the 10 selected TFs containing a GST tag at their C-terminus were expressed *in vitro* via the TnT SP6 High Yield Wheat

Germ Protein Expression System (Promega) according to the manufacturer's protocol. GST-labeled TFs were subsequently purified via glutathione sepharose 4B beads (GE Healthcare) and incubated overnight at 4°C on a rotator to enrich for the GST-fusion proteins. The beads were washed three times with a saline buffer (50 mM tris-HCl pH 8.0, 150 mM NaCl, 2mM EDTA, 0.1% NP-40, 10% glycerol) and boiled at 95°C for 5 min in protein loading buffer to release the fusion proteins. The protein-bead mixture was centrifuged at 6'000 rpm for 30 s to sediment the beads. The supernatant fraction was separated by SDS-PAGE in a 10% resolving gel (Tris-Glycine) and stained either with silver nitrate or with Coomassie blue (SimplyBlue Safestain, Invitrogen) for band visualization. Validation of *in vitro*-expressed proteins was performed by Western blot using 1:2000 diluted rabbit anti-GST primary antibodies (Cell signalling) and 1:5000 diluted goat anti-rabbit coupled to HRP secondary antibodies (Santa Cruz) on a nitrocellulose membrane. Validation of PPARγ and RXRα expression was performed again by Western blotting using 1:500 diluted rabbit anti-PPARγ (Santa Cruz, SC-7196) and 1:1000 diluted rabbit anti-RXRα (Santa Cruz, SC-553) primary antibodies, and using 1:1000 diluted PARP-1 as a nuclear control (Santa Cruz, SC-1561). Densitometric quantitation analyses were performed using the AlphaDigiDoc 1201 software (Alpha Innotech). (Semi)-quantitative Western blotting: for RXRα standard curve generation, *in vitro*-expressed RXRα was purified as described above. An aliquot was subjected to SRM for quantification (see below), which aided in the generation of an RXRα dilution series (0.31-20 ng). The corresponding RXRα standard curve samples and 3T3-L1 nuclear extracts harvested at different time points during differentiation were then Western blotted as described above..

For TF quantification: the 10 selected TFs were expressed *in vitro* via the TnT SP6 Wheat Germ High Yield Master Mix Minus Amino Acids (Promega) in their isotopically-labeled version. 18 amino acids were mixed to reach a 80µM final concentration and added to the reaction together with isotopically-labeled Arginine ($^{13}C,^{15}N$-Arg) and Lysine ($^{13}C,^{15}N$-Lys)(Cambridge Isotope Laboratories) to a 1mM final concentration prior to resorting to the conventional expression protocol. Depending on the analyzed TF, 3X reactions were utilized (or 4X for TFs with low expression yields). Subsequently, the purification and fractionation techniques described above were applied.

Assessment of protein losses during gel extraction

RXRα was *in vitro*-expressed and purified with glutathione sepharose 4B beads as described before. A series of dilutions (10 ng to 1.25 ng) was generated. One dilution set was used for validation by Western blotting using 1:1000 diluted rabbit anti-RXRα (Santa Cruz, SC-553) as described above. The second dilution set was loaded onto a 10% SDS-PAGE, stained with Coomassie blue after which bands corresponding to RXRα were excised and in gel-digested (as describe earlier). To assess protein losses during in-gel digestion, the resulting RXRα samples were further quantified by SRM.

Reference peptide tag incorporation

SH-Quant quantification strategy: The SH-Quant peptide (AADITSLYK)-coding sequence along with the segment of the 5' Gateway site comprised between the *Pvu*I and *Pvu*II restriction sites was synthesized by Geneart-Life Technologies. An *Spe*I restriction site was added at the 3' of the tag-coding sequence to simplify the insertion of tag variants. The SH-Quant insert was introduced in the *Pvu*I and *Pvu*II restriction sites at the 5' Gateway site of the pF3A-GST vector by double digestion with the two restriction enzymes and subsequently ligated to obtain the SH-pF3A-GST destination vector.

Flexi-Quant quantification strategy: The same cloning strategy as explained above was applied; here the SH-Quant peptide-coding sequence was replaced by a Flexi-Quant (TENLYFQGDISR)-coding sequence directly linked to a Strep-tag sequence (WSHPQFEK) at its 3' to ultimately obtain a Flex-pF3A-GST destination vector.

Reference peptide tag variants design for multiplexing

The additional tags 2 to 5 described in **Fig. 4** were designed by a first permutation of Ile residue 4 from the original SH-Quant tag sequence with amino acids of different hydrophobicity and size (G, A, V, F) to introduce molecular weight and/or retention time differences while not altering the fragmentation pattern. Additional tags 6 to 10 were generated by a second permutation conservative of the charge by replacing Asp residue 3 by Glu in the 5 sequences generated as above. Hydrophilic residues

Thr 5 and Tyr 8 which produce intense y" fragment ions were kept constant to maintain solubility and sensitivity. The *in vitro* expression vector SH-pF3A-GST was modified via site-directed mutagenesis (QuikChange Lightning, Stratagene) according to the manufacturer's protocol to generate the SH-quant peptide variant-incorporating vectors (SH_i-pF3A-GST). The I4G, I4A, I4V, I4F, D3E, D3E- I4G, D3E-I4A, D3E-I4V, and D3E-I4F substitutions were introduced into SH-pF3A-GST using the forward and reverse primers presented in **Supplementary Table 4**, and sequence-verified. Proteotypicity of the newly designed tags was confirmed with Blast-P against the mouse protein database (http://blast.ncbi.nlm.nih.gov).

In vitro-expressed TF identification

In vitro-expressed proteins were in-gel digested using trypsin as follows: the samples were reduced in 10 mM dithioerythritol (DTE) and alkylated in 55 mM iodoacetamide (IAA) before being dried. The samples were then incubated with 12.5 ug/ml of trypsin (in 50 mM ammonium bicarbonate, pH 7.5) overnight at 37°C. The resulting tryptic peptides were extracted from the gel slices, dried, and resuspended in 2% acetonitrile (ACN): 0.1% formic acid (FA) for LC-MS/MS analysis. A mass spectrometer (LTQ Orbitrap XL (Thermo Fisher Scientific) equipped with an ultraperformance LC (UPLC) system (nanoACQUITY; Waters) was used. Peptides were trapped in a custom-made precolumn (Magic C18 AQ stationary phase, 5 µm diameter, 200Å pore, 0.1 × 20 mm, Michrom Bioresource) and separated in a custom-made main column (Magic C18 AQ, 3 µm diameter, 100Å pore, 0.75 × 150 mm), using a run of 53 minutes and a gradient of H_2O:ACN:FA 98%:2%:0.1% (solvent A) and ACN:H_2O:FA 98%:1.9%:0.1% (solvent B). The gradient of the run was set at a flow rate of 250 nl/min as follows: 100% A for 3 min, 30% B within 36 min, 47% B within 14 min, 90% B within 5 min and held for 5 min and 100% A for 17 min. The MS/MS was operated in an information-dependent mode, in which each full MS analysis was followed by 10 MS/MS acquisitions during which the most abundant peptides were selected for collision-induced dissociation (CID) to generate tandem mass spectra. The normalized collision energies were set to 35% for CID. Data search was performed using Mascot software (v. 2.3; Matrix Science) and Proteome Discoverer (v.1.3; Thermo Fisher Scientific). The sequences were searched against a concatenated database consisting of a *Triticum aestivum* (wheat) database created from Uniprot database version 12.02 (4,684 sequences), the

sequences of the TF constructs expressed complemented with a set of common contaminant proteins sequences. Finally, the results were imported into Scaffold (v. 3.3; Proteome Software) for validation of protein identification, normalization, and comparison of spectral counts. Peptide identifications were accepted if they could be established at a probability of >95% as determined by the PeptideProphet algorithm[38]. Protein identifications were accepted if they were assigned at least two unique validated peptides, and could be established with at least 99% probability as determined by the ProteinProphet algorithm[39].

Proteotypic peptide selection

We initially turned to fragmentation spectra databases such as the NIST (http://www.nist.gov/nvl/) and PeptideAtlas[40] to select the ideal pool of peptide candidates for the design of our SRM-based assay, aiming to monitor nuclear PPARγ and RXRα abundance during terminal adipogenesis. However, only a limited amount of information regarding TF-specific peptides could be retrieved. We therefore had to devise a strategy that did not rely on the use of publicly available information to discover proteotypic peptides of TFs. To alleviate concerns regarding the detectability of TFs, we implemented three different bi-dimensional peptide fractionation pipelines on 3T3-L1 total nuclear protein extracts at day four of differentiation (a time point at which the adipogenic master regulators PPARγ and RXRα are known to be highly expressed[31, 41]) with the precise aim of discovering TF-specific peptide candidates that could be utilized in downstream quantitative analyses. However, only two peptides for each TF were retrieved in spite of the higher resolving power provided by the extensive fractionation and the consecutive higher sensitivity expected in mass spectrometry (**Supplementary Fig. 1**). Thus, despite recent increases in overall sensitivity[42], MS-based shotgun data-dependent methodologies are still lacking the combination of acquisition speed and sensitivity necessary to directly analyze low abundant proteins such as TFs within complex matrices derived from mammalian cells. The results of these preliminary experiments may also in part explain the scarcity of TF-specific peptides that we initially observed in databases. In response, we and others[14] are currently undertaking significant efforts to address this TF data paucity, but it is clear that current TF-specific peptide repositories are still far from complete. An alternative strategy based on the combined use of proteotypic peptide

prediction tools[43] and crude peptides synthesis[44] did also not improve the quality of our acquired data. In fact, a great majority of the 51 crude synthetic peptides selected were either hardly or not consistently detectable over the time course of the experiment (data not shown). We attribute this behavior to peptide solubility, stability, and ionisability issues and decided therefore not to continue with this strategy. However, we followed the steps indicated below to select the representative proteotypic peptides.

Peptide proteotypicity: candidate proteotypic peptides were selected using Pinpoint (v.1.1; Thermo Fischer Scientific) and their fragmentation spectra visually checked in Skyline[45] using the spectral library inspection function. The following rules were applied in Pinpoint and Skyline to the tryptic peptides detected to ensure that only high quality peptides were retained for the analysis.

Peptide uniqueness: as a general rule, the TF-specific peptides selected here were uniquely representative of the TF of interest. In some instances however, empirically observed best-flying candidates were shared among two or more TF isoforms or TF subfamily members and were included in the assay. Discarding such candidates for less-than-optimal ones could compromise detectability. Therefore, the precision of absolute measurements accomplished utilizing such non-unique TF-specific peptides may be affected to a certain extent by TF isoform(s) or subfamily counterpart(s). The following non-unique peptides were utilized in the multiplexed assay: ARID3a: peptide GLNLPTSITSAAFTLR is shared with ARID3b; NFIB: peptide EDFVLTVTGK is shared with NFIA, and GIPLESTDGER is shared with NFIA and NFIC; PIAS3: peptide VSELQVLLGFARG is shared with PIAS2; RXRα: peptide VLTELVSK is shared with RXRβ and ILEAELAVEPK is shared with RXRα; SMAD2: peptide VETPVLPPVLVPR is shared with SMAD3.

Peptide detectability: only peptides spanning 7-25 amino-acids and falling within a mass range of 700 to 3,000 Da were used. N– or C–terminal peptides which could be degraded or modified in the protein of interest were excluded. The sequences immediately flanking the N– and C–terminal peptides of the target sequence were also excluded from the evaluation (the sequence linking the SH-Quant tag to the N–terminus and the sequence linking the C–terminal peptide and the GST-tag sequence).

Peptide digestability: as the whole methodology is peptide-centric and thus dependent on the completeness and specificity of the digestion step[46], special care

was taken to ensure that only end-product peptides would be selected for the assay. This was especially true for the SH-Quant tag construction which contains a [KK] dibasic sequence at its C–terminal part which could lead to miscleavages. The SH-Quant tag configuration was however retained as systematic attempts to detect miscleavage products originating from the [KK] dibasic sequence using either accurate mass inclusion mass spectrometry or targeted SRM detection were negative (i.e. only the SH-Quant tag end-product was detected). For all the other peptides, only fully tryptic peptides with no miscleavage other than [KP] and [RP] were considered. N– or C– terminal dibasic residues ([KK], [KR], [RK], [RR]) as well as surrounding acidic residues (E, D) in position P'1 and P'2 potentially leading to miscleavages were excluded.

Peptide stability: peptide sequences containing oxidation-sensitive residues, methionine only in this case, were excluded. These peptides which usually ionize well were however kept as secondary choices in case no usable sequences would remain after the selection[47]. Peptides containing potential imide forming residues, [NG], potential deamidating residue, [DP], or N–terminal glutamine (Q) were excluded as well.

Peptide selection: A first manual ranking of the proteotypic peptides was performed using peptide intensities extracted from the MS1 stage with the intention of selecting the best "flyers". This ranking was complemented by a manual evaluation of the tandem MS spectral quality aimed at not excluding low intensity MS1 peaks with intense selective fragment ions (TIC MS2 in **Supplementary Table 2**). Specifically, tandem MS spectra presenting high intensity singly charged y'' ions above the m/z of the parent ion were included in the second screen performed on the triple quadrupole instrument (TSQ Vantage, Thermo Fisher Scientific). In a preliminary SRM run, all selected peptides were targeted. The final best-responding candidates were selected by performing a pilot SRM analysis targeting the heavy labeled versions of the peptides selected in the previous screen spiked in a nuclear extract digest prepared in the same way as for the real experiment. A minimum of two peptides per TF was finally selected for the multiplex assay.

Nuclear extract sample preparation and spiking of in vitro-expressed heavy TF and SH-Quant reference tags

Samples containing *in vitro*-expressed heavy TF were run on separate gel lanes allowing for maximal protein recovery. Gel bands were excised and split in four fractions of equal size, one of which was promptly utilized and the remaining ones were stored for future needs. Typically, expression levels of 3x to 4x were sufficient for 4 different multiplex assays. All *in vitro*-expressed heavy TFs were initially assessed for heavy isotope incorporation levels during a targeted SRM run. If required, this initial step can be utilized for external heavy TF quantitation using the SH-Quant tag, hence providing estimates on (spike-in) loads for a multiplex assay. This external quantitation is also used for precise adjustment of the light and heavy SH-Quant tag ratio, which in turn is fundamental for accurate quantitation using the multiplex approach. The basis for choosing the right quantities that need to be spiked in the case of TFs was determined by the necessity of achieving a good signal for the heavy labeled SH-Quant tags in the nuclear matrix (multiplex). For nuclear extracts (NE), 400 µg (quantified using the BCA assay) of total NE protein (differentiation times: Day 0 to Day 6) were separated and fractionated by SDS-PAGE using 10% Tris-Glycine gels. Gel bands were excised around the migration height of the proteins of interest: 50-60 kDa for RXRα and PPARγ, 30-40 kDa and 45-70 kDa in the multiplexed analysis (encompassing therefore all 10 TFs).

Gel fractions of NE and heavy TFs were subsequently pooled for reduction and alkylation followed trypsin digestion, as described above. The accurately quantified SH-Quant tags (JPT Peptide Technologies) were added to each tube (200 fmol per tag) immediately after re-suspension in the digestion buffer. Following trypsinization, peptides were extracted from gel fractions and accurately aliquoted into four equal volumes (technical replicates). Samples were then dried using a speed vacuum and stored at –20°C prior to analysis.

Liquid chromatography-mass spectrometry (LC-MS) analysis

Dried aliquots (theoretical concentration: 100 µg) were re-suspended in 20 µl LC-MS loading solvent (2% acetonitrile, 0.1% formic acid) yielding a final, theoretical concentration of nuclear extract of 5 µg/µl. Following re-suspension, samples were allowed to settle for 1h to increase overall peptide solubility prior to analysis. Typically, 5 µl of sample (theoretical total: 25 µg) was loaded and captured on a home-made capillary pre-column (C18; 3 µm-200Å; 2 cm × 250 µm) prior to analytical LC separation (nanoACQUITY UPLC, Waters). Samples were separated

using a 90-minute biphasic gradient starting from 100% A solvent (2% acetonitrile, 0.1% formic acid) to 90% B solvent (100% acetonitrile, 0.1% formic acid) on a Nikkyo (Nikkyo Technologies) nano-column (C18; 3 µm-100 Å; 15 cm length and 100 µm inner diameter, at a flow of 1 µl/min). The gradient was followed by a wash for 8 min at 90% solvent B and column re-equilibration of 12 min at 100% solvent A.

Selected reaction monitoring (SRM) of TFs

All samples were analyzed on a TSQ-Vantage triple-quadrupole mass spectrometer (Thermo Fisher Scientific). A 0.7 FWHM resolution window for both Q1 and Q3 was set for parent- and product-ion isolation. Fragmentation of parent-ions was performed in Q2 at 1.5 mTorr, using collision energies calculated with the Pinpoint software (v1.1). Parent-ion selection was set for fully digested peptides on the doubly charged ion for both SH-Quant tags and target proteins. Generally, singly charged peptide fragment ions ranging from y''_3 to y''_{n-1} were preferably monitored, unless otherwise stated. A complete list of all monitored transitions is provided in **Supplementary Data 2**. A parent-ion retention time target window of 2 min to quantify RXRα or PPARγ and 5 min for the multiplexed quantification of ten TFs was set for Q1 during a scheduled SRM run. A total of respectively 88 and 76 transitions were monitored for RXRα and PPARγ in the single protein quantification strategy, whereas a total of 492 transitions were typically monitored during a multiplex run. A cycle time of 0.5 s and 2 s was used for single and multiplexed SRM runs respectively. A minimum dwell time of 10 ms or more was set for both single and multiplexed SRM assays. Samples were analyzed as three biological and three technical replicates (n=3) in RXRα and PPARγ single protein quantification assays. One biological replicate was analyzed in 6 technical replicates as a proof-of-concept of the multiplex SRM assay.

Evaluation of heavy TF isotope incorporation and miscleavage during SRM

All *in vitro*-expressed heavy TFs were assessed for light isotope (K & R) misincorporation prior to SRM quantification. The amounts of isotope misincorporation (usually ≤ 5%) were corrected in the final calculations of absolute endogenous concentrations. In addition, all SH-Quant tags were monitored for trypsin miscleavage during a dedicated LC-MS run targeting the SH-Quant tag and its potential miscleavage products to assess the extent of digestion completeness.

Data analysis and absolute quantification of TFs

Calculation of absolute levels of TFs was performed using a two-step procedure as outlined in **Fig. 1b**. All data analyses were carried out using Pinpoint software (v.1.1). Peptide identification and peak area integration of all SH-Quant tags and targeted peptides as well as their transitions were manually verified in Pinpoint. Ten individual standard curves were established for all SH-Quant tags using a concentration range of six values recorded in three technical replicates (**Supplementary Fig. 14**). The performance-based definitions of lower limits of quantification (LLOQ) were applied as defined by the FDA (Guidance for Industry: Bioanalytical Method Validation: *http://www.fda.gov/cder/guidance/index.htm*). Briefly, the threshold for the limit of detection (LOD) was set at 20% of the CV, whereas the LLOQ was set at 15% of the CV using three technical replicates for single- and six technical replicates for multiplex quantification approaches. An example of the LLOQ found for an SH-Tag is provided in **Supplementary Fig. 15**. An analysis of variance (ANOVA) coupled with a Neuwman-Keuls post-hoc test was applied for calculating the differences in endogenous levels of TFs between different differentiation time points, where the level of $P<0.05$ was considered as statistically significant. The calculation steps to derive absolute copy numbers per cell (nucleus) are presented in **Supplementary Data 1**.

SRM assay validation using accurately quantified proteins

Recombinant human α-synuclein (SYN, ^{15}N heavy isotope, MW: 14,625 Da) and histidyl-tRNA synthetase (HARS, MW: 59,849 Da, ProspecBio, Israel) were accurately quantified by UV spectrophotometry using the protein's molar extinction coefficient at 280nm and by amino acid analysis (AAA). The AAA was performed twice on two dilutions of the original protein solution at the Functional Genomics Center, Zurich (FGCZ) and gave a precision of ± 5%. Protein amounts corresponding to 5,000 (SYN = 3.64 ng, HARS = 14.3 ng) and 50,000 (SYN = 36.4 ng, HARS = 143 ng) copies were spiked into nuclear extract (D0, 400 µg) and analyzed by SDS-PAGE. SYN and HARS proteins were cut at their corresponding gel migration height as determined by reference proteins. In parallel, we retrieved the respective human ORFs (HARS, Thermo Fisher; SYN, in-house) and cloned them into our expression vector to enable the production of the respective heavy isotope-labeled proteins and thus to allow for a similar experimental workflow as the one used to quantify RXRα

and PPARγ. Gel bands were pooled and digested with the heavy SH-tagged protein and the accurately quantified SH-tags as outlined in the multiplex approach. SRM quantitation of spiked human (quantified by AAA) SYN and HARS peptides was achieved in a single SRM run using an 80 min LC gradient. A total of ten peptides were selected for HARS protein of which seven peptides share sequence homology with the mouse HARS protein, two peptides are proteotypic for human- and one peptide for mouse (endogenous) HARS. For SYN, two proteotypic peptides were monitored. Despite the high sequence homology between human and mouse SYN, endogenous levels of SYN could be accounted for by monitoring the ratio between the light (mouse endogenous) and the heavy ^{15}N labeled peptides against the *in vitro*-expressed heavy SYN protein labeled at the terminal lysine. Samples were analyzed in three technical replicates.

Motif Analysis

Regions enriched with H3K27Ac as determined by Mikkelsen *et al.,*[29] were scanned with FIMO[48] allowing for a *P* value of 10^{-3}. Each match in those regions was then scored for log-likelihood given by the position weight matrix of the PPARγ motif from the JASPAR CORE database[49], assuming uniform background.

PPARγ ChIP-seq data reanalysis

ChIP-seq data from Nielsen *et al*. ([26]) was processed as in Raghav *et al.*[31].

Quantitative model of genome-wide TF DNA binding

The estimation of the PPARγ DNA binding profile is based on equilibrium thermodynamics and is described in detail in the **Supplementary Note**.

References

1. Simicevic, J. & Deplancke, B. DNA-centered approaches to characterize regulatory protein-DNA interaction complexes. *Mol. Biosyst.* **6**, 462-468 (2010).

2. Kim, H.D., Shay, T., O'Shea, E.K. & Regev, A. Transcriptional Regulatory Circuits: Predicting Numbers from Alphabets. *Science* **325**, 429-432 (2009).

3. Bussemaker, H.J., Foat, B.C. & Ward, L.D. Predictive Modeling of Genome-Wide mRNA Expression: From Modules to Molecules. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 329-347 (2007).

4. Segal, E. & Widom, J. From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat. Rev. Genet.* **10**, 443-456 (2009).

5. Stormo, G.D. & Zhao, Y. Determining the specificity of protein–DNA interactions. *Nat. Rev. Genet.* **11**, 751-760 (2010).

6. Biggin, Mark D. Animal Transcription Networks as Highly Connected, Quantitative Continua. *Dev. Cell* **21**, 611-626 (2011).

7. Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A. & Luscombe, N.M. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10**, 252-263 (2009).

8. Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W. & Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6940-6945 (2003).

9. Picotti, P., Bodenmiller, B., Mueller, L.N., Domon, B. & Aebersold, R. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell* **138**, 795-806 (2009).

10. Picotti, P. et al. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* **494**, 266-270 (2013).

11. Kuster, B., Schirle, M., Mallick, P. & Aebersold, R. Scoring proteomes with proteotypic peptide probes. *Nat. Rev. Mol. Cell Biol.* **6**, 577-583 (2005).

12. Brun, V. et al. Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol. Cell. Proteomics* **6**, 2139-2149 (2007).

13. Hanke, S., Besir, H., Oesterhelt, D. & Mann, M. Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J. Proteome Res.* **7**, 1118-1130 (2008).

14. Stergachis, A.B., Maclean, B., Lee, K., Stamatoyannopoulos, J.A. & Maccoss, M.J. Rapid empirical discovery of optimal peptides for targeted proteomics. *Nat. Meth.* **8**, 1041-1043 (2011).

15. Zeiler, M., Straube, W.L., Lundberg, E., Uhlen, M. & Mann, M. A protein epitope signature Tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines. *Mol. Cell. Proteomics* **11**, O111.009613 (2012).

16. Pratt, J.M. et al. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat. Prot.* **1**, 1029-1043 (2006).

17. Holzmann, J., Pichler, P., Madalinski, M., Kurzbauer, R. & Mechtler, K. Stoichiometry determination of the MP1-p14 complex using a novel and cost-efficient method to produce an equimolar mixture of standard peptides. *Anal. Chem.* **81**, 10254-10261 (2009).

18. Singh, S., Springer, M., Steen, J., Kirschner, M.W. & Steen, H. FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. *J Proteome Res.* **8**, 2201-2210 (2009).

19. Wepf, A., Glatter, T., Schmidt, A., Aebersold, R. & Gstaiger, M. Quantitative interaction proteomics using mass spectrometry. *Nat. Meth.* **6**, 203-205 (2009).

20. Proc, J.L. et al. A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *J. Proteome Res.* **9**, 5422-5437 (2010).

21. Kuhn, E. et al. Inter-laboratory evaluation of automated, multiplexed peptide immunoaffinity enrichment coupled to multiple reaction monitoring mass spectrometry for quantifying proteins in plasma. *Mol. Cell. Proteomics* **11**, M111.013854 (2012).

22. Ahmed, M., Neville, M.J., Edelmann, M.J., Kessler, B.M. & Karpe, F. Proteomic analysis of human adipose tissue after rosiglitazone treatment shows coordinated changes to promote glucose uptake. *Obesity* **18**, 27-34 (2010).

23. Molina, H. et al. Temporal profiling of the adipocyte proteome during differentiation using a five-plex SILAC based strategy. *J. Proteome Res.* **8**, 48-58 (2009).

24. Mirzaei, H., McBee, J.K., Watts, J. & Aebersold, R. Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. *Mol. Cell. Proteomics* **7**, 813-823 (2008).

25. Dupuis, A., Hennekinne, J.A., Garin, J. & Brun, V. Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. *Proteomics* **8**, 4633-4636 (2008).

26. Nielsen, R. et al. Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev.* **22**, 2953-2967 (2008).

27. Kaplan, T. et al. Quantitative Models of the Mechanisms That Control Genome-Wide Patterns of Transcription Factor Binding during Early <italic>Drosophila</italic> Development. *PLoS Genet.* **7**, e1001290 (2011).

28. Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* **451**, 535-540 (2008).

29. Mikkelsen, T.S. et al. Comparative epigenomic analysis of murine and human adipogenesis. *Cell* **143**, 156-169 (2010).

30. Rey, G. et al. Genome-Wide and Phase-Specific DNA-Binding Rhythms of BMAL1 Control Circadian Output Functions in Mouse Liver. *PLoS Biol.* **9**, e1000595 (2011).

31. Raghav, Sunil K. et al. Integrative Genomics Identifies the Corepressor SMRT as a Gatekeeper of Adipogenesis through the Transcription Factors C/EBPβ and KAISO. *Mol. Cell* **46**, 335-350 (2012).

32. Siersbaek, R. et al. Extensive chromatin remodelling and establishment of transcription factor /`hotspots/' during early adipogenesis. *EMBO J.* **30**, 1459-1472 (2011).

33. Lamesch, P. et al. hORFeome v3.1: A resource of human open reading frames representing over 10,000 human genes. *Genomics* **89**, 307-315 (2007).

34. Hens, K. et al. Automated protein-DNA interaction screening of Drosophila regulatory elements. *Nat. Meth.* **8**, 1065-1070 (2011).

35. Whiteaker, J.R. et al. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat. Biotech.* **29**, 625-634 (2011).

36. Farnham, P.J. Insights from genomic profiling of transcription factors. *Nat. Rev. Genet.* **10**, 605-616 (2009).

37. John, S. et al. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat. Genet.* **43**, 264-268 (2011).

38. Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383-5392 (2002).

39. Nesvizhskii, A.I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646-4658 (2003).

40. Desiere, F. et al. Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol.* **6**, R9 (2005).

41. Rosen, E.D. & MacDougald, O.A. Adipocyte differentiation from the inside out. *Nat. Rev. Mol. Cell Biol.* **7**, 885-896 (2006).

42. Thakur, S.S. et al. Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol. Cell. Proteomics* **10**, M110 003699 (2011).

43. Prakash, A. et al. Expediting the development of targeted SRM assays: using data from shotgun proteomics to automate method development. *J. Prot. Res.* **8**, 2733-2739 (2009).

44. Picotti, P. et al. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* **7**, 43-46 (2010).

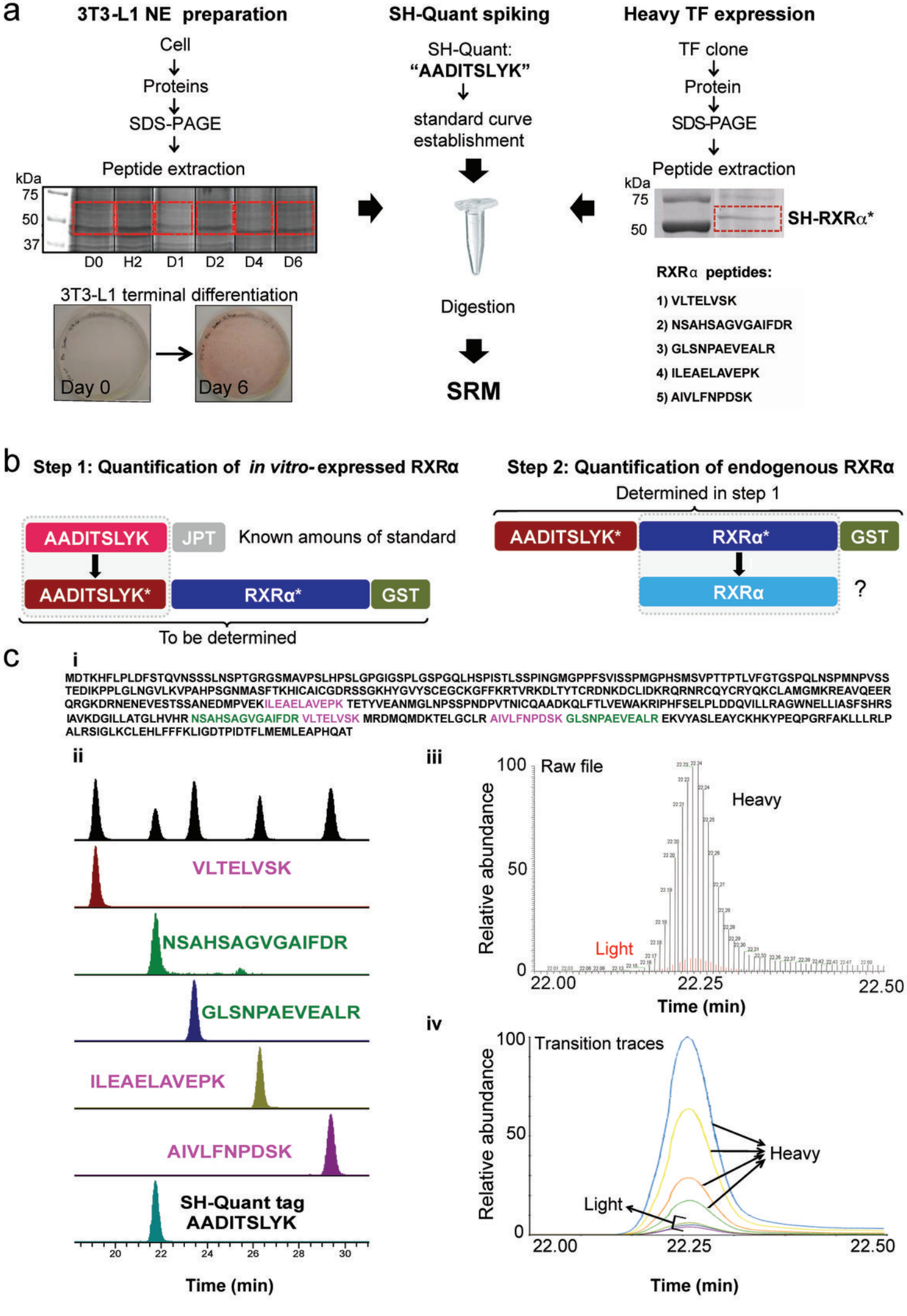
45. MacLean, B. et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968 (2010).

46. Brownridge, P. & Beynon, R.J. The importance of the digest: Proteolysis and absolute quantification in proteomics. *Methods* **54**, 351-360 (2011).

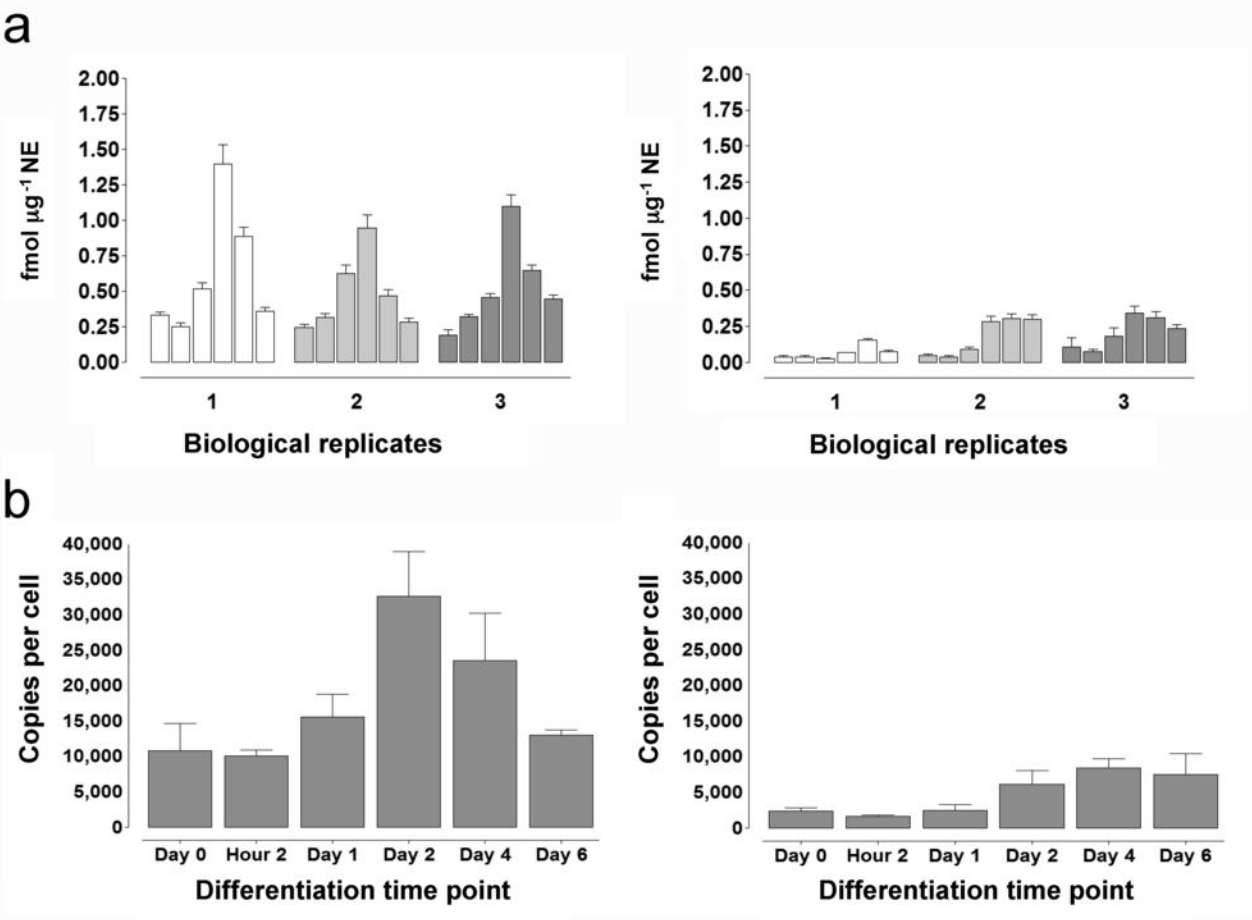
47. Jaquinod, M. et al. Mass spectrometry-based absolute protein quantification: PSAQ™ strategy makes use of “noncanonical” proteotypic peptides. *Proteomics* **12**, 1217-1221 (2012).

48. Grant, C.E., Bailey, T.L. & Noble, W.S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011).

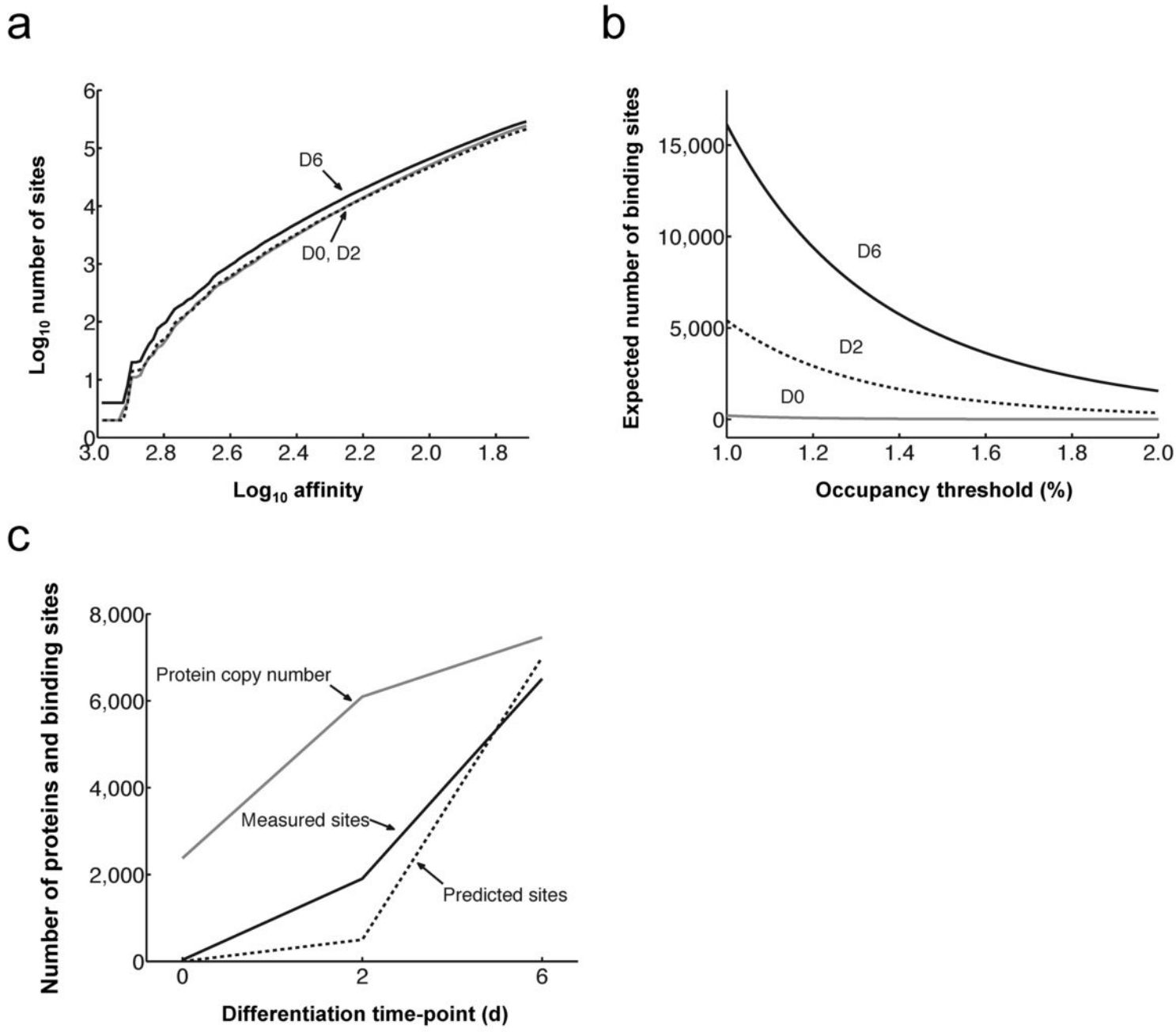
49. Vlieghe, D. et al. A new generation of JASPAR, the open-access repository for transcription factor binding site profiles. *Nucleic Acids Res.* **34**, D95-97 (2006).


a
3T3-L1 NE preparation
Cell
Proteins
SDS-PAGE
Peptide extraction
kDa
75
50
37
D0
H2
D1
D2
D4
D6
3T3-L1 terminal differentiation
Day 0
Day 6
SH-Quant spiking
SH-Quant:
"AADITSLYK"
standard curve establishment
Digestion
SRM
Heavy TF expression
TF clone
Protein
SDS-PAGE
Peptide extraction
kDa
75
50
SH-RXRα*
RXRα peptides:
1) VLTELVSK
2) NSAHSAGVGAIFDR
3) GLSNPAEVEALR
4) ILEAELAVEPK
5) AIVLFNPDSK
b
Step 1: Quantification of in vitro-expressed RXRα
AADITSLYK
JPT
Known amouns of standard
AADITSLYK*
RXRα*
GST
To be determined
Step 2: Quantification of endogenous RXRα
Determined in step 1
AADITSLYK*
RXRα*
GST
RXRα
?
c
i
MDTKHFLPLDFSTQVNSSSLNSPTGRGSMAVPSLHPSLGPGIGSPLGSPGQLHSPISTLSSPINGMGPPFSVISSPMGPHSMSVPTTPTLVFGTGSPQLNSPMNPVSS
TEDIKPPLGLNGVLKVPAHPSGNMASFTKHICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDNKDCLIDKRQRNRCQYCRYQKCLAMGMKREAVQEER
QRGKDRNENEVESTSSANEDMPVEKILEAELAVEPK TETYVEANMGLNPSSPNDPVTNICQAADKQLFTLVEWAKRIPHFSELPLDDQVILLRAGWNELLIASFSHRS
IAVKDGILLATGLHVHR NSAHSAGVGAIFDR VLTELVSK MRDMQMDKTELGCLR AIVLFNPDSK GLSNPAEVEALR EKVYASLEAYCKHKYPEQPGRFAKLLLRLP
ALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEAPHQAT
ii
VLTELVSK
NSAHSAGVGAIFDR
GLSNPAEVEALR
ILEAELAVEPK
AIVLFNPDSK
SH-Quant tag
AADITSLYK
20
22
24
26
28
30
Time (min)
iii
Raw file
Heavy
Light
Relative abundance
100
50
0
22.00
22.25
22.50
Time (min)
iv
Transition traces
Heavy
Light
Relative abundance
100
50
0
22.00
22.25
22.50
Time (min)

Simicevic *et al*. Figure 2



Simicevic *et al.* Figure 3

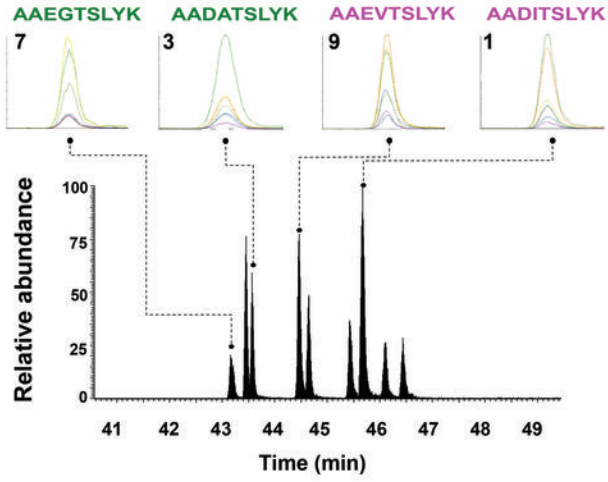


Simicevic *et al*. Figure 4

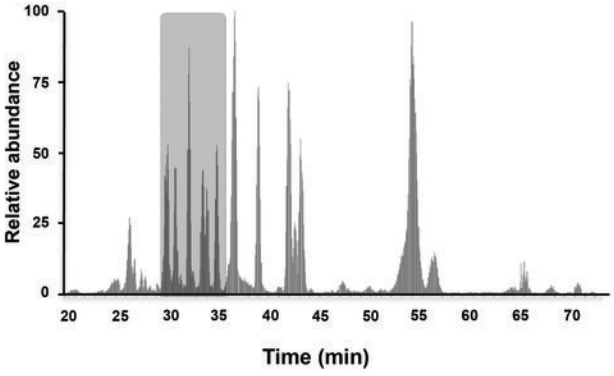
a

no.	TF	Tag sequence	Parent Mass m/z [M+2H]$^{2+}$	RT min	Transitions [M+H]$^{+}$ y8	y7	y6	y5	y4	y3	y2
1	RXRα	AADITSLYK	491.26	46.3	910	839	724	611	510	423	310
2	Nfib	AADGTSLYK	463.24	43.3	854	783	668	611	510	423	310
3	Pias3	AADATSLYK	470.25	43.8	868	797	682	611	510	423	310
4	Pias4	AADVTSLYK	484.25	45.2	896	825	710	611	510	423	310
5	Fosl2	AADFTSLYK	508.25	47.3	944	873	758	611	510	423	310
6	Rarg	AAEITSLYK	498.27	46.2	924	853	724	611	510	423	310
7	PPARγ	AAEGTSLYK	470.24	43.2	868	797	668	611	510	423	310
8	Arid3a	AAEATSLYK	477.25	43.7	882	811	682	611	510	423	310
9	Nr2c1	AAEVTSLYK	491.26	45.0	910	839	710	611	510	423	310
10	Smad2	AAEFTSLYK	515.26	47.1	958	887	758	611	510	423	310

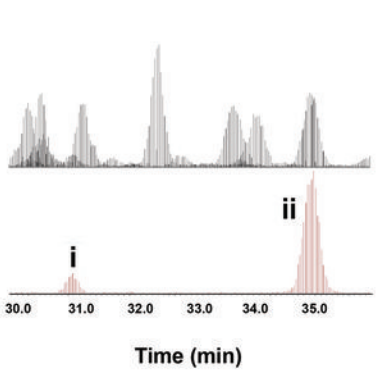
b



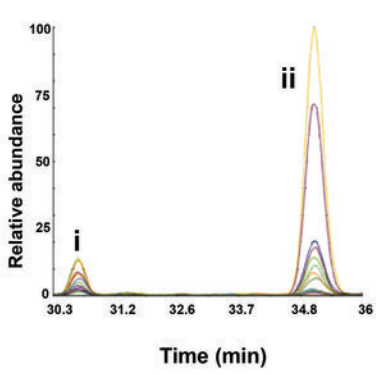
c



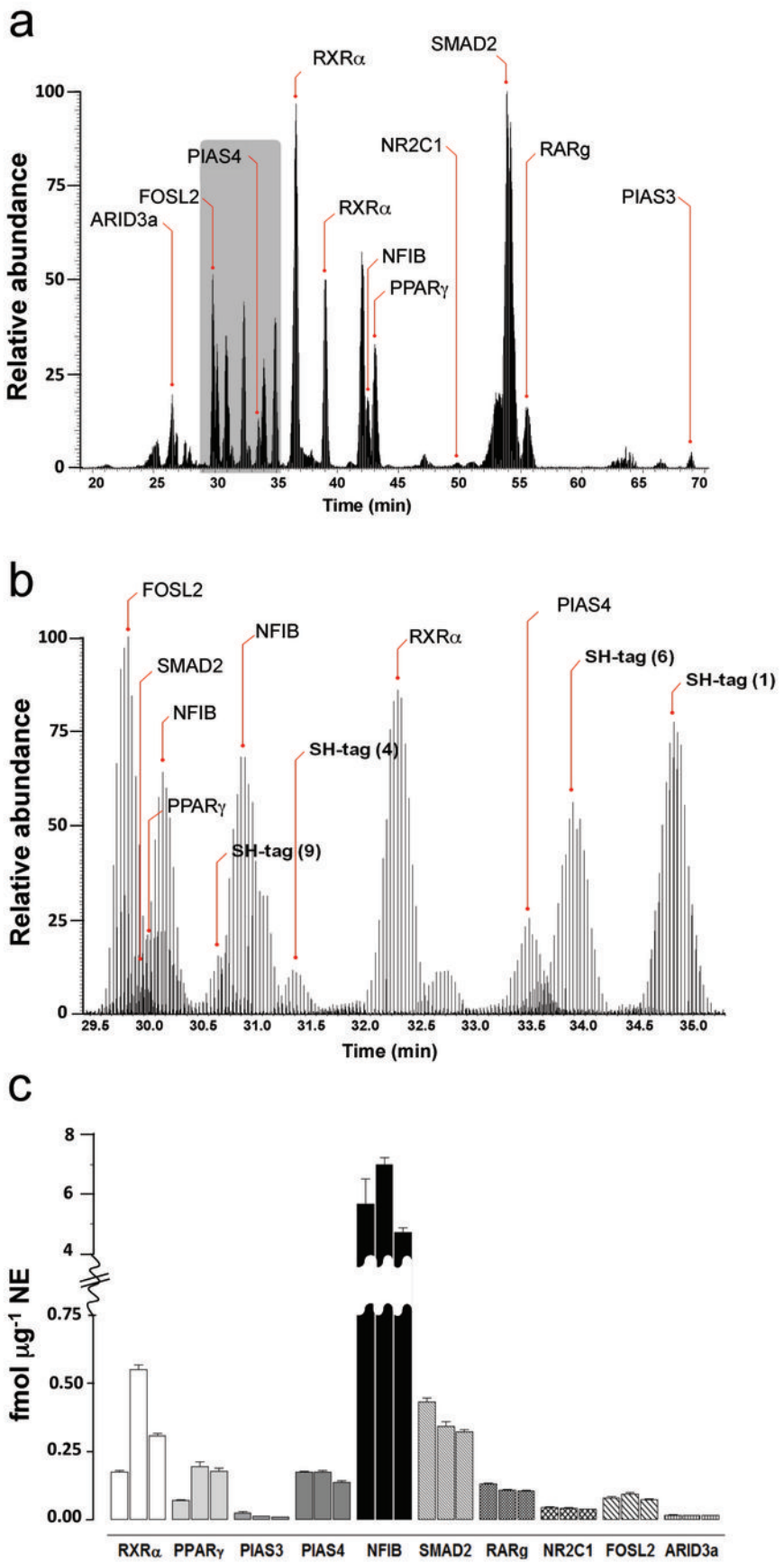
d



e

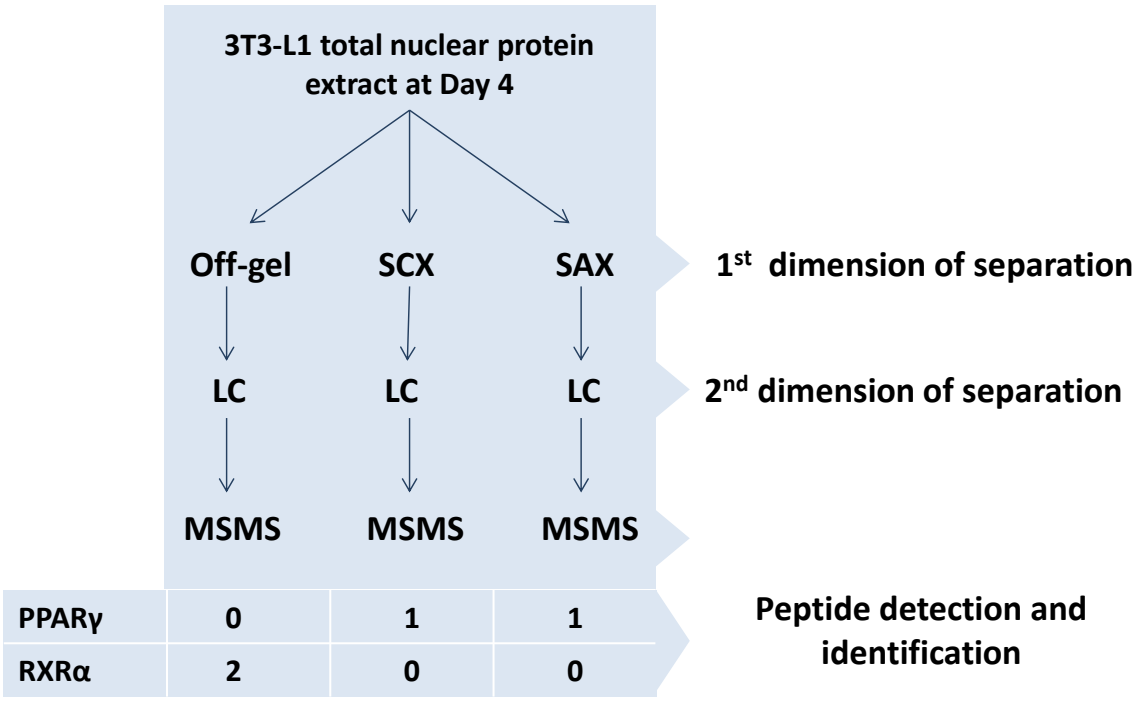


Simicevic *et al*. Figure 5



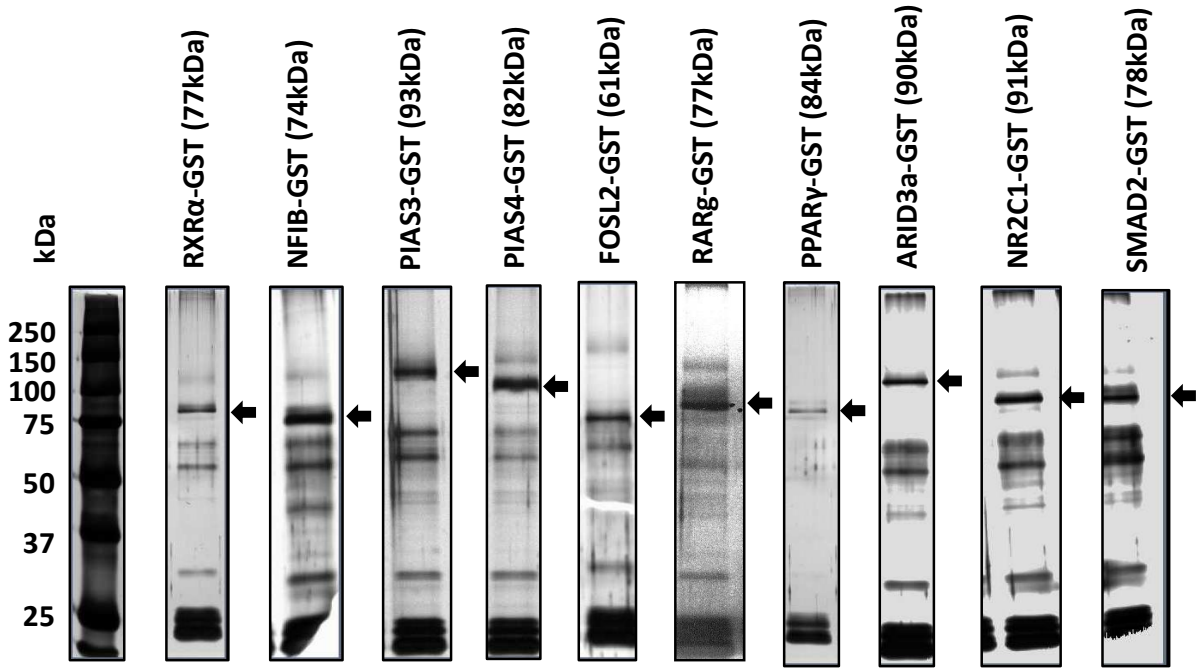
Supplementary Figures

Supplementary Figure 1.



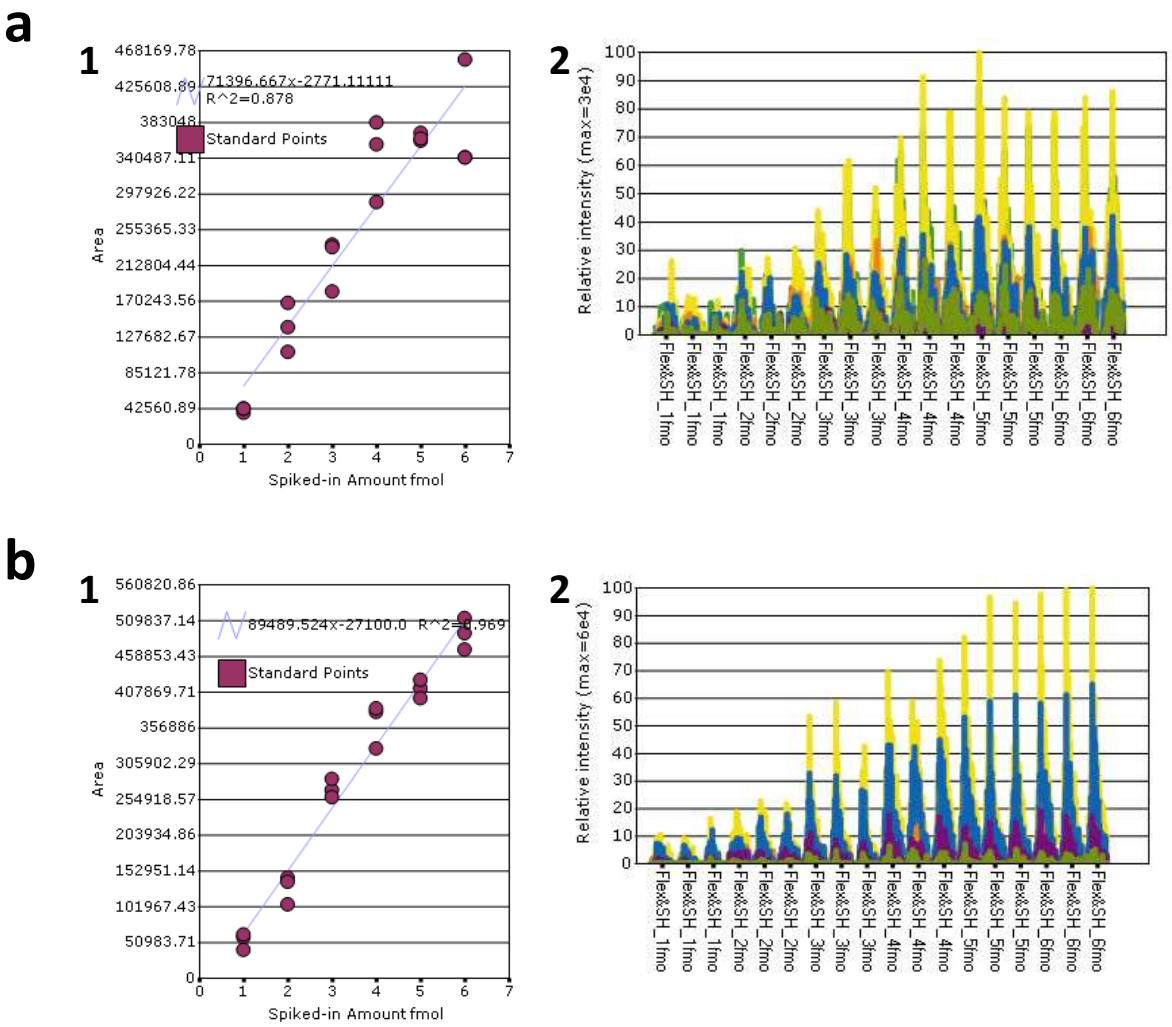
Supplementary Figure 1. Diagram representing the 3 *shotgun* MS pipelines tested utilizing 3T3-L1 total nuclear protein extract at Day 4 of differentiation.
Each pipeline employs a different peptide separation technique in the first dimension: SAX (strong anion exchange, 6 fractions), SCX (strong cation exchange, 20 fractions), and separation of peptides by their isoelectric point using an "Off-gel" electrophoretic system (24 fractions). The peptides were subsequently submitted to LC-MSMS for identification. The two adipogenic master regulators PPARγ and RXRα were specifically targeted. A maximum of only two peptides per TF were ultimately observed.

Supplementary Figure 2.



Supplementary Figure 2. 10 *in vitro*-expressed GST-tagged TFs.
The 10 *in vitro*-expressed TFs were run on a denaturing SDS-PAGE gel stained with silver nitrate. The molecular weights of the TF constructs (molecular weight of the TF + molecular weight of the GST-tag (26 kDa)) are presented in parentheses and their respective bands are indicated by an arrow. Peptide detection and protein identification for each TF construct was performed by in-gel digestion and LC-MS analysis.

Supplementary Figure 3.



Supplementary Figure 3. Comparison of the performance of the FLEXIQuant and SH-Quant peptide tags in SRM.

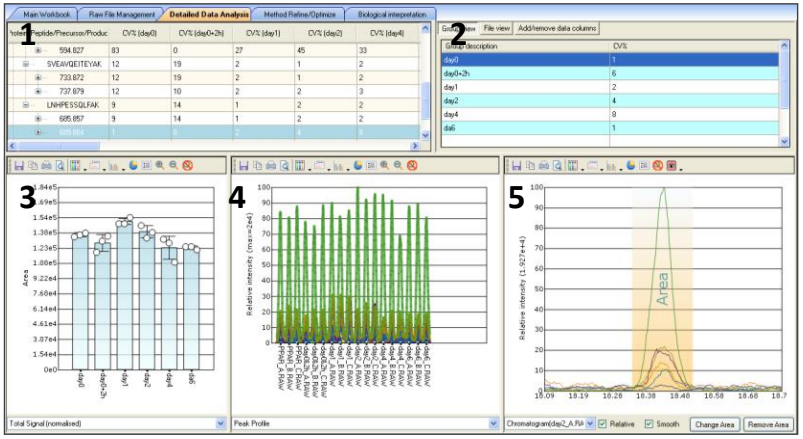
(**a**) FLEXIQuant tag: Subpanel 1 shows the calculated standard curve regression line for the FLEXIQuant peptide (R^2=0.878). Subpanel 2 shows peak profiles of different spikes of the FLEXIQuant peptide (1-6 fmol) in three technical replicates.

(**b**) SH-Quant tag: Subpanel 1 shows the calculated standard curve regression line for the SH-Quant peptide (R^2=0.969). Subpanel 2 shows peak profiles of different spikes of the SH-Quant peptide (1-6 fmol) in three technical replicates.

Supplementary Figure 4.

a

PPARγ peptide: LNHPESSQLFAK

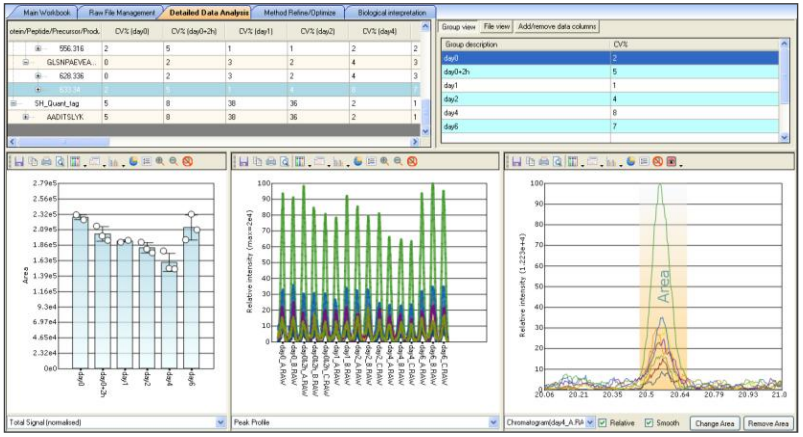


b

PPAR_A.RAW	13.12 fmol
PPAR_B.RAW	13.11 fmol
PPAR_C.RAW	13.09 fmol
day0&2h_A.RAW	13.25 fmol
day0&2h_B.RAW	13.16 fmol
day0&2h_C.RAW	13.11 fmol
day1_A.RAW	13.01 fmol
day1_B.RAW	13.01 fmol
day1_C.RAW	12.96 fmol
day2_A.RAW	13.03 fmol
day2_B.RAW	13.13 fmol
day2_C.RAW	13.08 fmol
day4_A.RAW	13.14 fmol
day4_B.RAW	13.17 fmol
day4_C.RAW	13.33 fmol
day6_A.RAW	13.21 fmol
day6_B.RAW	13.2 fmol
day6_C.RAW	13.23 fmol

c

RXRα peptide: GLSNPAEALR



d

day0_A.RAW	8.25 fmol
day0_B.RAW	8.11 fmol
day0&2h_A.RAW	7.93 fmol
day0&2h_B.RAW	7.64 fmol
day0&2h_C.RAW	7.5 fmol
day1_A.RAW	7.45 fmol
day1_B.RAW	7.51 fmol
day2_A.RAW	7.45 fmol
day2_B.RAW	7.25 fmol
day2_C.RAW	7.14 fmol
day4_A.RAW	7.2 fmol
day4_B.RAW	6.69 fmol
day4_C.RAW	6.66 fmol
day6_A.RAW	7.53 fmol
day6_B.RAW	8.27 fmol
day6_C.RAW	7.83 fmol

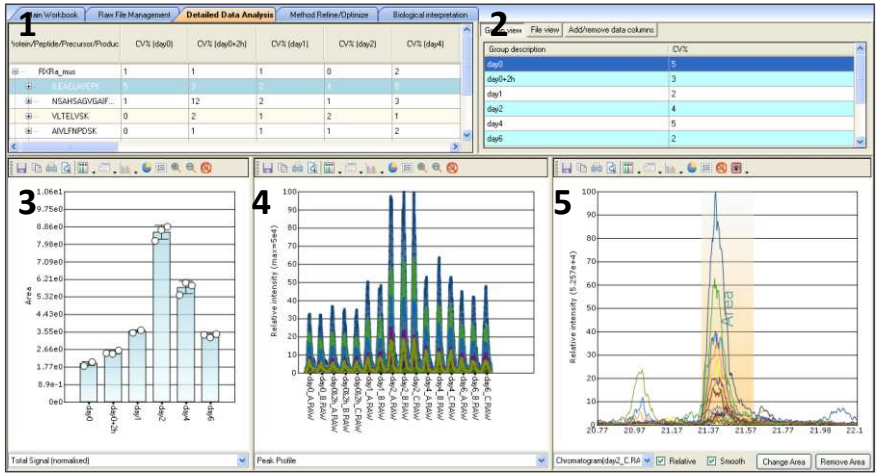
Supplementary Figure 4. Evaluation of heavy PPARγ "LNHPESSQLFAK" and RXRα "GLSNPAEALR" peptides in 3T3-L1 nuclear extract.

Pinpoint screen prints, showing an example of the evaluation of heavy PPARγ (**a**) (LNHPESSQLFAK) and RXRα (**c**) (GLSNPAEALR) peptides, monitored at Day 0 through Day 6. The Pinpoint window provides a summary of the evaluation of the peptide of interest. Subpanel 1 shows an overview of all peptides monitored. Subpanel 2 shows the coefficients of variation per peptide at the different time points. Subpanel 3 shows the total normalized signal of all technical replicates (n=3). Subpanel 4 shows the peak profiles of all technical replicates. Subpanel 5 shows the chromatogram and calculated peak area of peptide transitions. Panel legend and numbering applies to panel (**c**) as well. Panels **b** and **d** show the calculated values of the PPARγ and RXRα heavy peptides spiked into nuclear extracts at Day 0 through Day 6.

Supplementary Figure 5.

a

RXRα peptide: ILEAELAVEPK

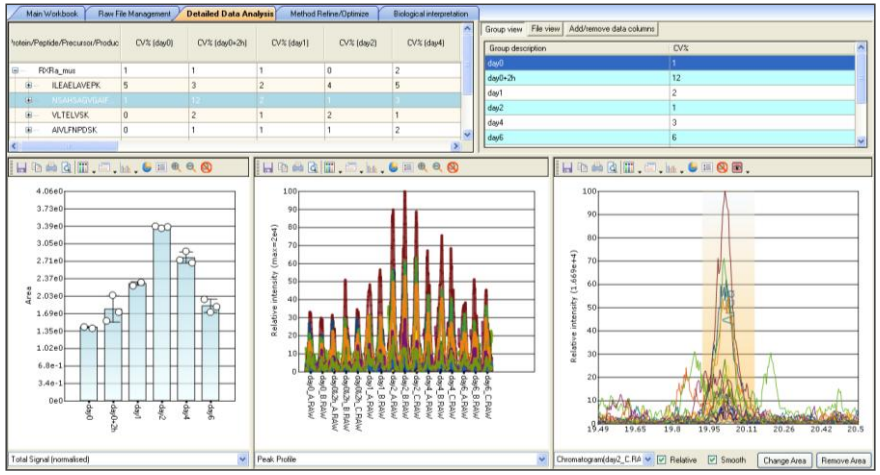


b

day0_A.RAW	15.13 fmol
day0_B.RAW	16.76 fmol
day0&2h_A.RAW	18.87 fmol
day0&2h_B.RAW	18.4 fmol
day0&2h_C.RAW	19.75 fmol
day1_A.RAW	25.39 fmol
day1_B.RAW	26.28 fmol
day2_A.RAW	62 fmol
day2_B.RAW	66.29 fmol
day2_C.RAW	67.37 fmol
day4_A.RAW	35.7 fmol
day4_B.RAW	39.8 fmol
day4_C.RAW	38.73 fmol
day6_A.RAW	27.63 fmol
day6_B.RAW	26.37 fmol
day6_C.RAW	27.84 fmol

c

RXRα peptide: NSAHSAGVGAIF

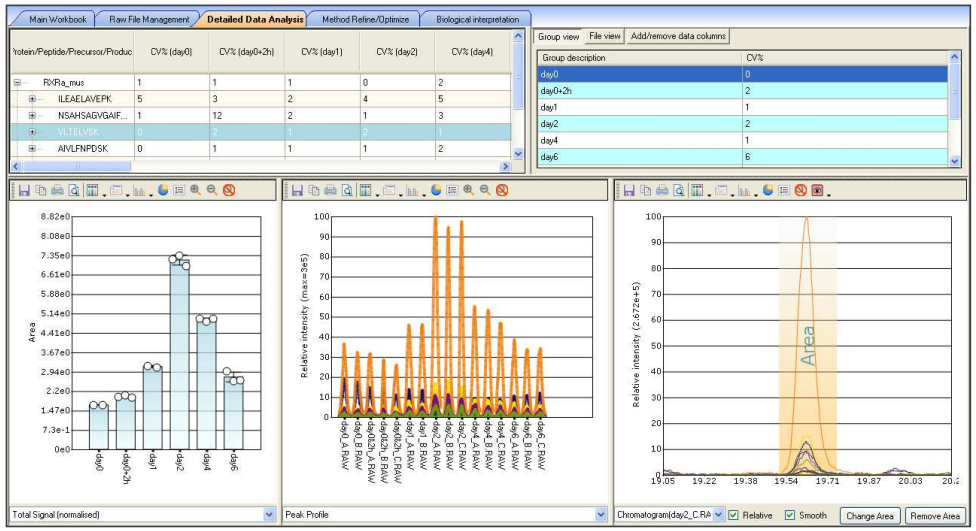


d

day0_A.RAW	11.8 fmol
day0_B.RAW	11.58 fmol
day0&2h_A.RAW	11.73 fmol
day0&2h_B.RAW	15.57 fmol
day0&2h_C.RAW	13.03 fmol
day1_A.RAW	16.16 fmol
day1_B.RAW	16.66 fmol
day2_A.RAW	25.74 fmol
day2_B.RAW	25.42 fmol
day2_C.RAW	25.62 fmol
day4_A.RAW	17.98 fmol
day4_B.RAW	19.1 fmol
day4_C.RAW	17.67 fmol
day6_A.RAW	16 fmol
day6_B.RAW	13.94 fmol
day6_C.RAW	14.77 fmol

RXRα peptide: VLTELVSK

e

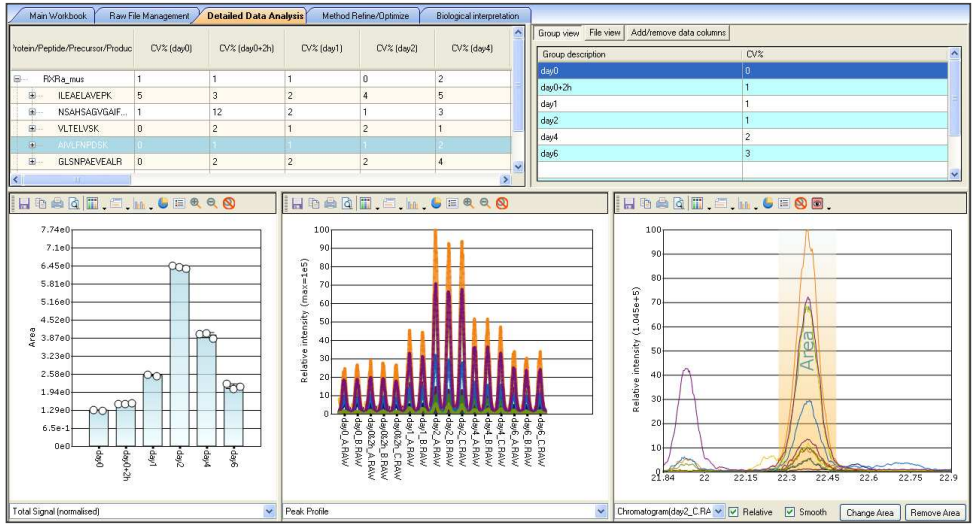


f

day0_A.RAW	13.9 fmol
day0_B.RAW	13.83 fmol
day0&2h_A.RAW	15.13 fmol
day0&2h_B.RAW	15.65 fmol
day0&2h_C.RAW	14.91 fmol
day1_A.RAW	22.97 fmol
day1_B.RAW	22.4 fmol
day2_A.RAW	54.79 fmol
day2_B.RAW	55.85 fmol
day2_C.RAW	52.89 fmol
day4_A.RAW	32.66 fmol
day4_B.RAW	31.95 fmol
day4_C.RAW	32.58 fmol
day6_A.RAW	24.15 fmol
day6_B.RAW	21.11 fmol
day6_C.RAW	21.37 fmol

RXRα peptide: AIVLFNPDSK

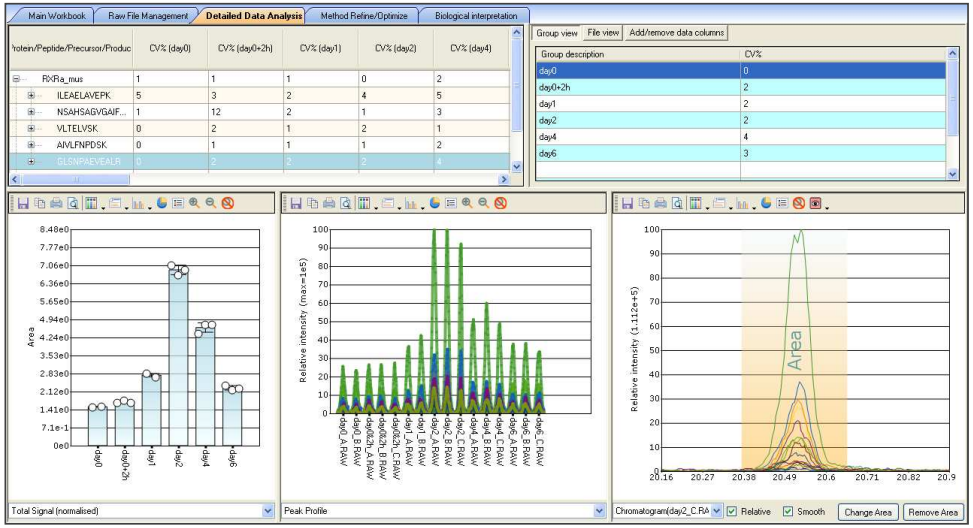
g



h

day0_A.RAW	10.6 fmol
day0_B.RAW	10.5 fmol
day0&2h_A.RAW	11.48 fmol
day0&2h_B.RAW	11.42 fmol
day0&2h_C.RAW	11.79 fmol
day1_A.RAW	18.52 fmol
day1_B.RAW	18.1 fmol
day2_A.RAW	49.04 fmol
day2_B.RAW	48.78 fmol
day2_C.RAW	48.34 fmol
day4_A.RAW	26.41 fmol
day4_B.RAW	26.62 fmol
day4_C.RAW	25.45 fmol
day6_A.RAW	18.08 fmol
day6_B.RAW	16.76 fmol
day6_C.RAW	17.26 fmol

i **RXRα peptide: GLSNPAEVEALR**



j

day0_A.RAW	12.5 fmol
day0_B.RAW	12.59 fmol
day0&2h_A.RAW	12.79 fmol
day0&2h_B.RAW	13.44 fmol
day0&2h_C.RAW	12.86 fmol
day1_A.RAW	20.44 fmol
day1_B.RAW	19.49 fmol
day2_A.RAW	53.68 fmol
day2_B.RAW	50.87 fmol
day2_C.RAW	52.48 fmol
day4_A.RAW	28.89 fmol
day4_B.RAW	31.34 fmol
day4_C.RAW	31.18 fmol
day6_A.RAW	19.16 fmol
day6_B.RAW	17.71 fmol
day6_C.RAW	18.31 fmol

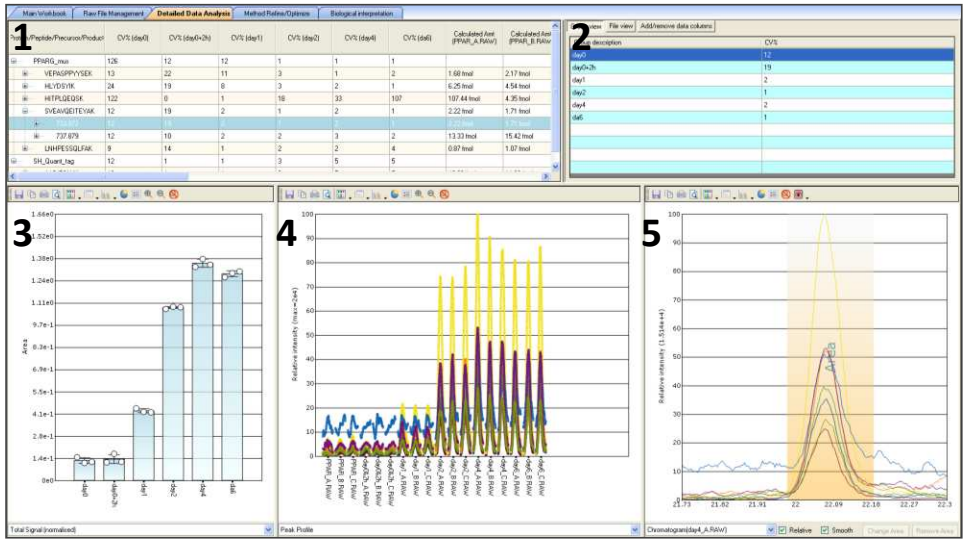
Supplementary Figure 5. SRM monitoring of the five RXRα peptides in 3T3-L1 nuclear extract.

(**a,c,e,g,i**) Screen prints of all five RXRα peptides monitored by SRM. The Pinpoint window provides a summary of the evaluation of the peptide of interest. Please see Supplementary Fig. 3 for a description of the subpanels. (**b,d,f,h,j**) Calculated amounts of endogenous levels found in nuclear extracts.

Supplementary Figure 6.

a

PPARγ peptide: SVEAQEITEYAK

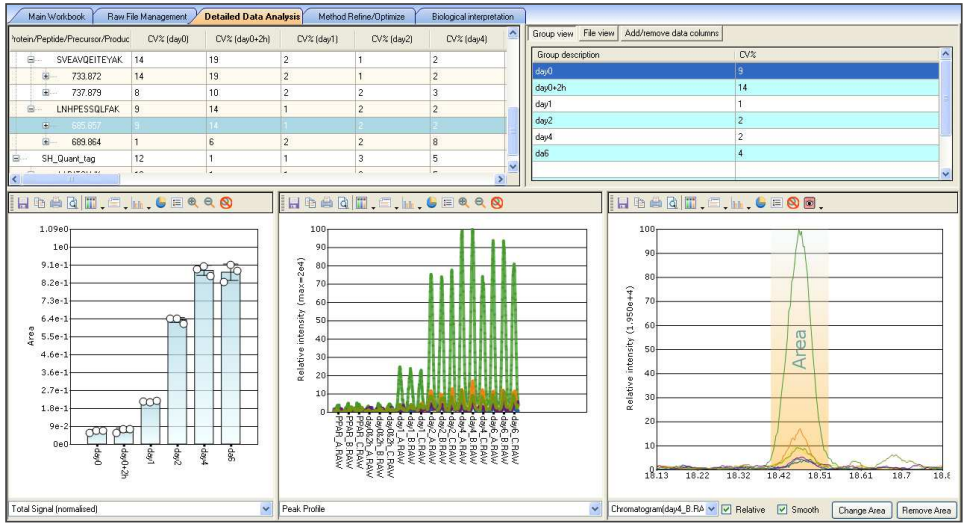


b

PPAR_A.RAW	2.22 fmol
PPAR_B.RAW	1.71 fmol
PPAR_C.RAW	1.8 fmol
day0&2h_A.RAW	1.58 fmol
day0&2h_B.RAW	2.32 fmol
day0&2h_C.RAW	1.62 fmol
day1_A.RAW	5.06 fmol
day1_B.RAW	4.84 fmol
day1_C.RAW	4.78 fmol
day2_A.RAW	14.61 fmol
day2_B.RAW	14.79 fmol
day2_C.RAW	14.73 fmol
day4_A.RAW	16.46 fmol
day4_B.RAW	17.11 fmol
day4_C.RAW	16.66 fmol
day6_A.RAW	15.88 fmol
day6_B.RAW	16.19 fmol
day6_C.RAW	16.33 fmol

c

PPARγ peptide: LNHPESSQLFAK

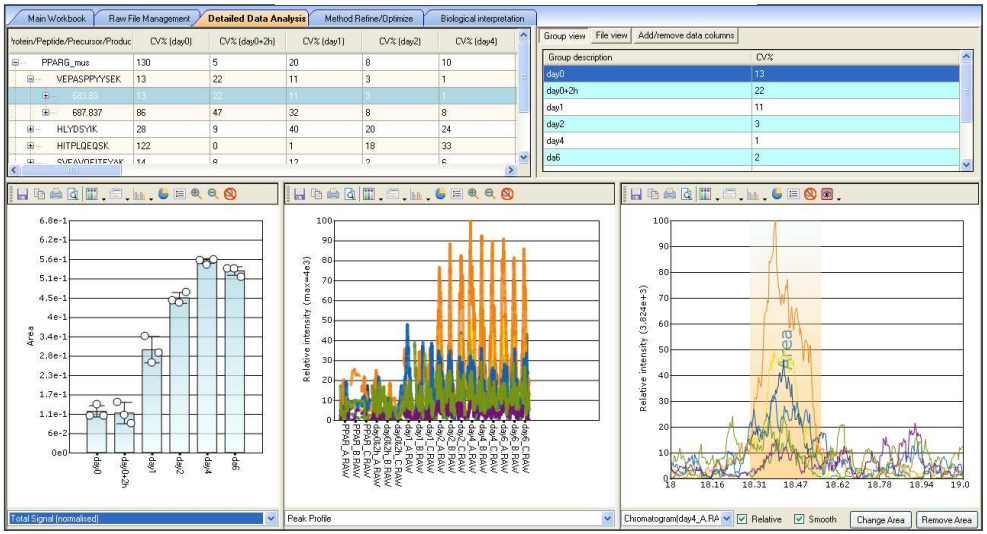


d

PPAR_A.RAW	0.87 fmol
PPAR_B.RAW	1.07 fmol
PPAR_C.RAW	1.07 fmol
day0&2h_A.RAW	0.79 fmol
day0&2h_B.RAW	1.08 fmol
day0&2h_C.RAW	1.07 fmol
day1_A.RAW	2.43 fmol
day1_B.RAW	2.4 fmol
day1_C.RAW	2.48 fmol
day2_A.RAW	8.72 fmol
day2_B.RAW	8.72 fmol
day2_C.RAW	8.38 fmol
day4_A.RAW	10.99 fmol
day4_B.RAW	11.19 fmol
day4_C.RAW	10.57 fmol
day6_A.RAW	10.35 fmol
day6_B.RAW	11.4 fmol
day6_C.RAW	11.03 fmol

PPARγ peptide: VEPASPPYYSEK

e

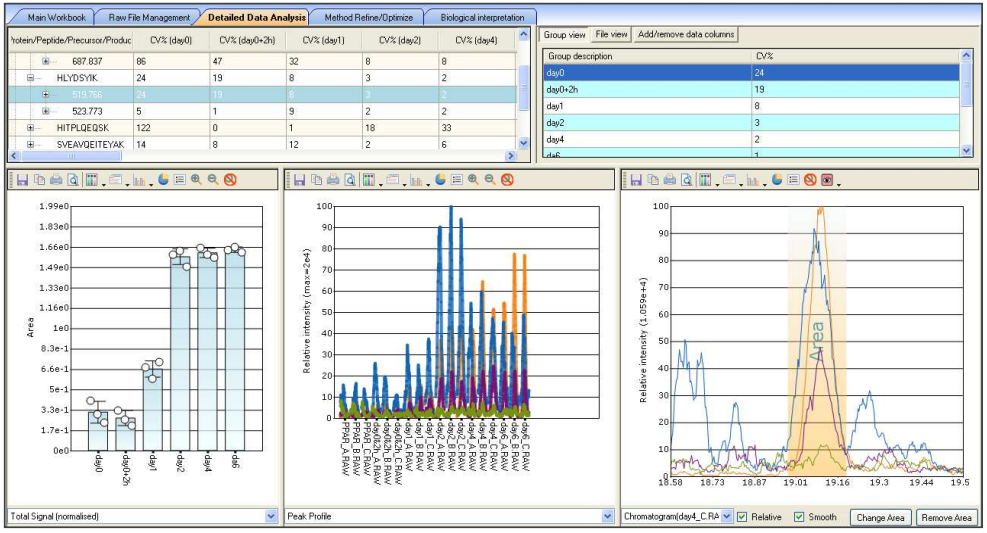


f

PPAR_A.RAW	1.68 fmol
PPAR_B.RAW	2.17 fmol
PPAR_C.RAW	1.67 fmol
day0&2h_A.RAW	2.06 fmol
day0&2h_B.RAW	1.53 fmol
day0&2h_C.RAW	1.2 fmol
day1_A.RAW	3.86 fmol
day1_B.RAW	2.99 fmol
day1_C.RAW	3.3 fmol
day2_A.RAW	6.08 fmol
day2_B.RAW	5.99 fmol
day2_C.RAW	6.42 fmol
day4_A.RAW	6.97 fmol
day4_B.RAW	6.82 fmol
day4_C.RAW	6.99 fmol
day6_A.RAW	6.75 fmol
day6_B.RAW	6.75 fmol
day6_C.RAW	6.44 fmol

PPARγ peptide: HLYDSYIK

g



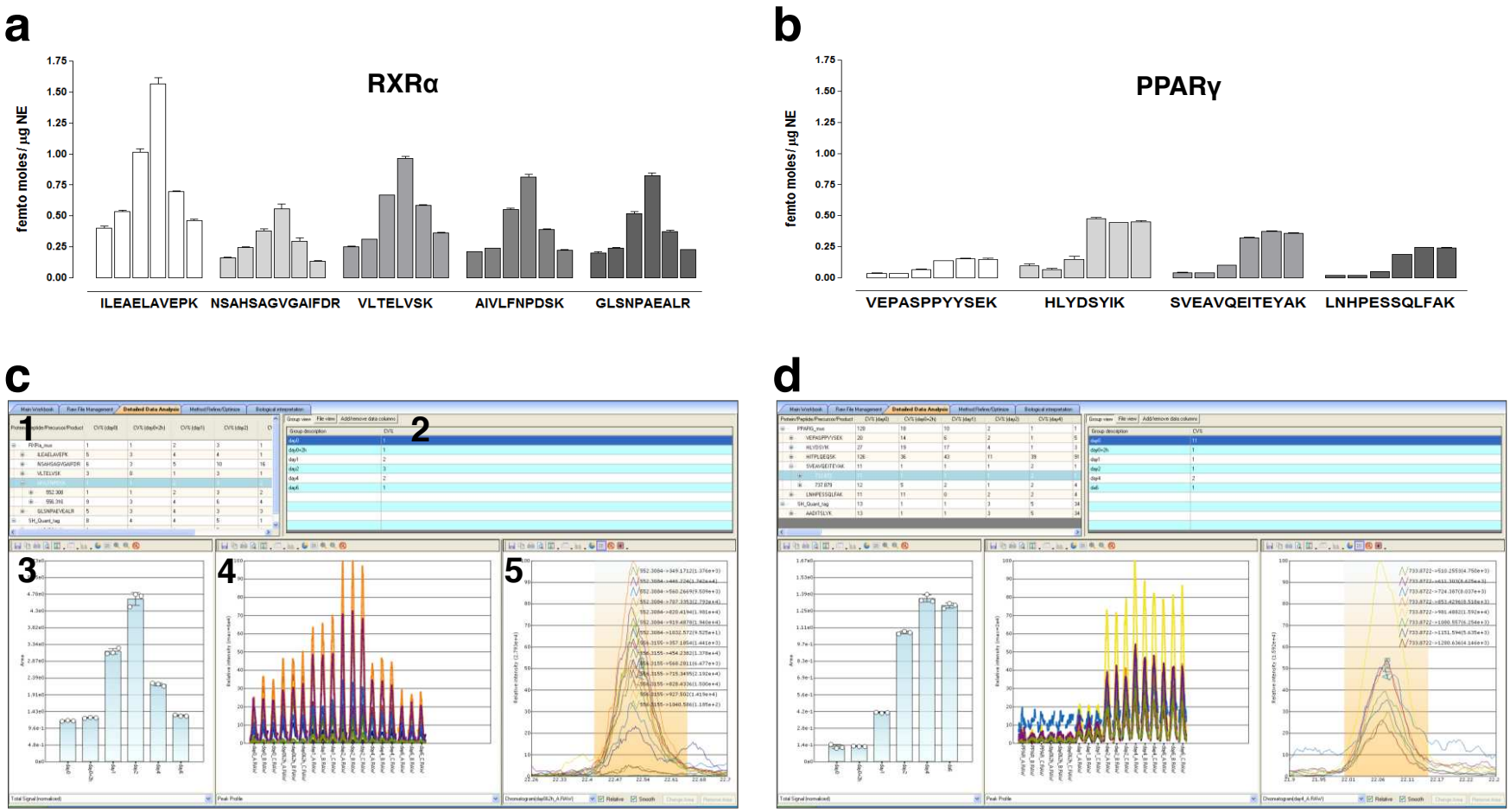
h

PPAR_A.RAW	6.25 fmol
PPAR_B.RAW	4.54 fmol
PPAR_C.RAW	3.53 fmol
day0&2h_A.RAW	4.56 fmol
day0&2h_B.RAW	3.56 fmol
day0&2h_C.RAW	2.86 fmol
day1_A.RAW	7.66 fmol
day1_B.RAW	6.65 fmol
day1_C.RAW	8.15 fmol
day2_A.RAW	21.88 fmol
day2_B.RAW	22.29 fmol
day2_C.RAW	20.53 fmol
day4_A.RAW	20.48 fmol
day4_B.RAW	19.84 fmol
day4_C.RAW	19.52 fmol
day6_A.RAW	20.5 fmol
day6_B.RAW	20.8 fmol
day6_C.RAW	20.24 fmol

Supplementary Figure 6. SRM monitoring of the four PPARγ peptides in 3T3-L1 nuclear extract.

(**a,c,e,g**) Screen prints of all four PPARγ peptides monitored by SRM. The Pinpoint window provides a summary of the evaluation of the peptide of interest. Please see Supplementary Fig. 3 for a description of the subpanels. (**b,d,f,h**) Calculated amounts of endogenous levels found in nuclear extracts.

Supplementary Figure 7.



e

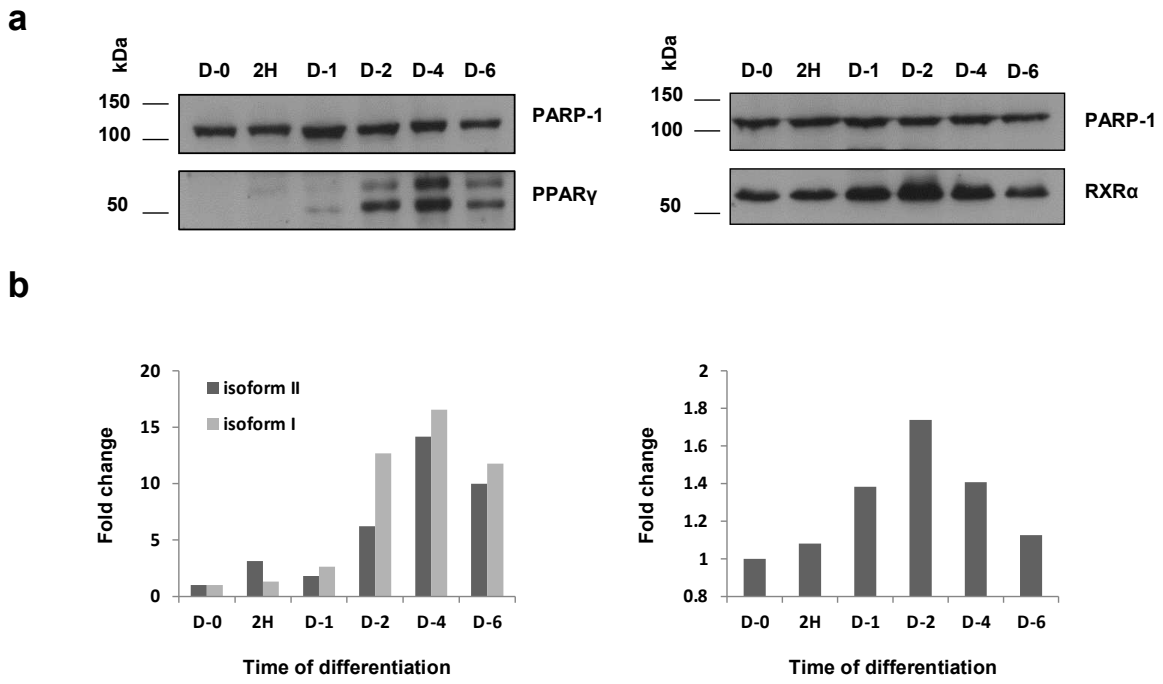
RXRα sequence	Parent Mass m/z $[M+2H]^{2+}$	CV of peptides analyzed at day 0 to day 6					
		day 0	day 0 + 2h	day 1	day 2	day 4	day 6
ILEAELAVEPK	606.34	5%	3%	4%	4%	1%	3%
NSAHSAGVGAIFDR	701.34	6%	3%	5%	10%	16%	5%
VLTELVSK	444.77	3%	<1%	1%	3%	1%	2%
AIVLFNPDSK	552.3	1%	1%	2%	3%	2%	1%
GLSNPAEALR	628.33	5%	3%	4%	3%	3%	2%

f

PPARγ sequence	Parent Mass m/z $[M+2H]^{2+}$	CV of peptides analyzed at day 0 to day 6					
		day 0	day 0 + 2h	day 1	day 2	day 4	day 6
VEPASPPYYSEK	683.83	20%	14%	6%	2%	1%	5%
HLYDSYIK	519.76	27%	19%	17%	4%	1%	3%
SVEAVQEITEYAK	733.87	11%	1%	1%	1%	2%	1%
LNHPESSQLAK	685.85	11%	11%	<1%	2%	2%	4%

Supplementary Figure 7. Calculated coefficient of variation of all RXRα and PPARγ peptides monitored by SRM.
(**a-b**) Summary of all RXRα and PPARγ peptides monitored by SRM. Values are the mean ± SD of one biological sample analyzed in three technical replicates. The Pinpoint window provides a summary of the results as well as the calculated CV of RXRα (**c**) and PPARγ (**d**) peptides respectively. Please see Supplementary Fig. 3 for a description of the subpanels. Tables (**e-f**) summarize the calculated CV values for all RXRα (n=5) and PPARγ (n=4) peptides monitored by SRM.

Supplementary Figure 8.

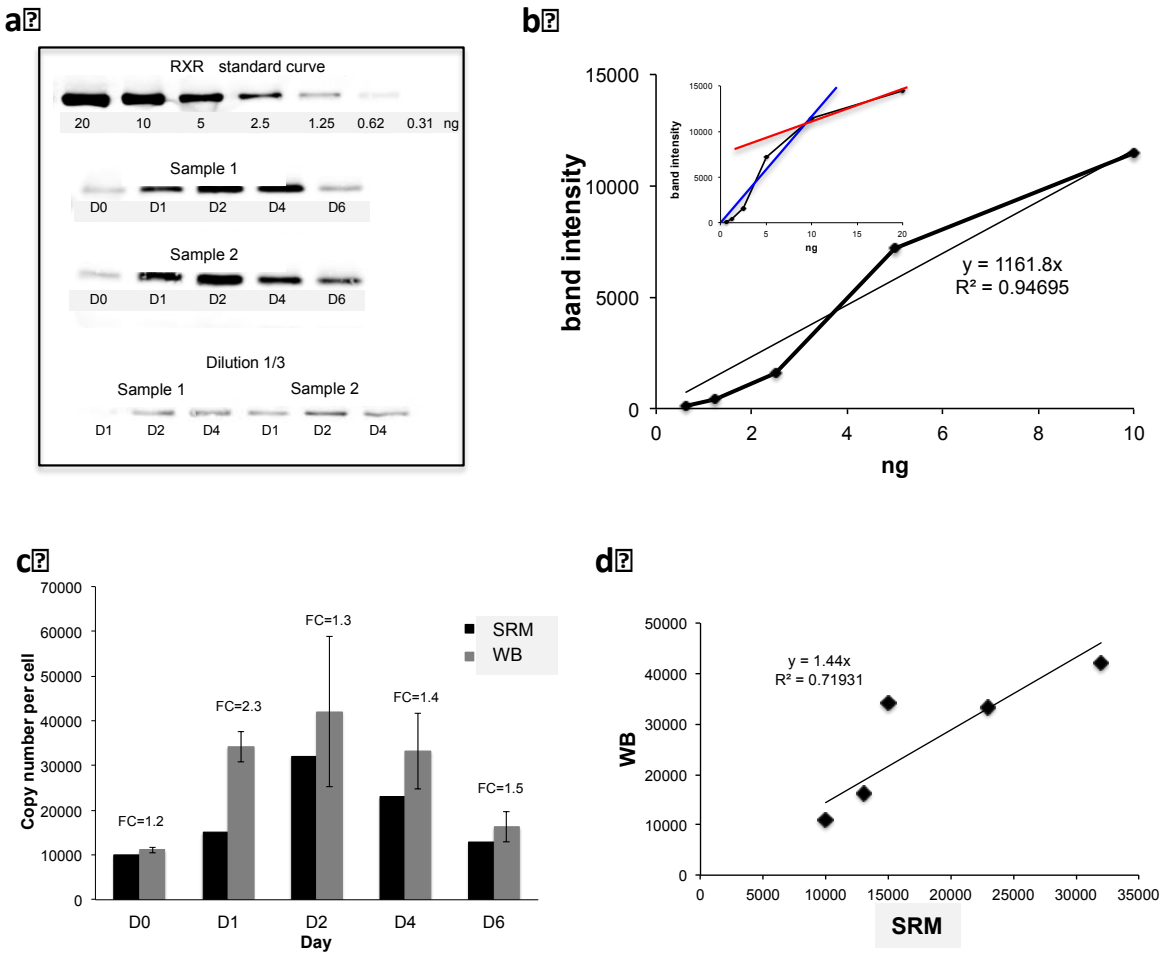


Supplementary Figure 8. Immunoblotting analysis of PPARγ and RXRα expression levels found in 3T3-L1 nuclear extracts.

(**a**) Validation of nuclear levels of PPARγ (left panel: isoform II on top at 57.6 kDa, isoform I on the bottom at 54.5 kDa) and RXRα (right panel: 51,2 kDa) expression by Western blotting using TF-specific primary antibodies in 3T3-L1 cells during six time points of terminal differentiation. PARP-1 was used as a nuclear loading control.

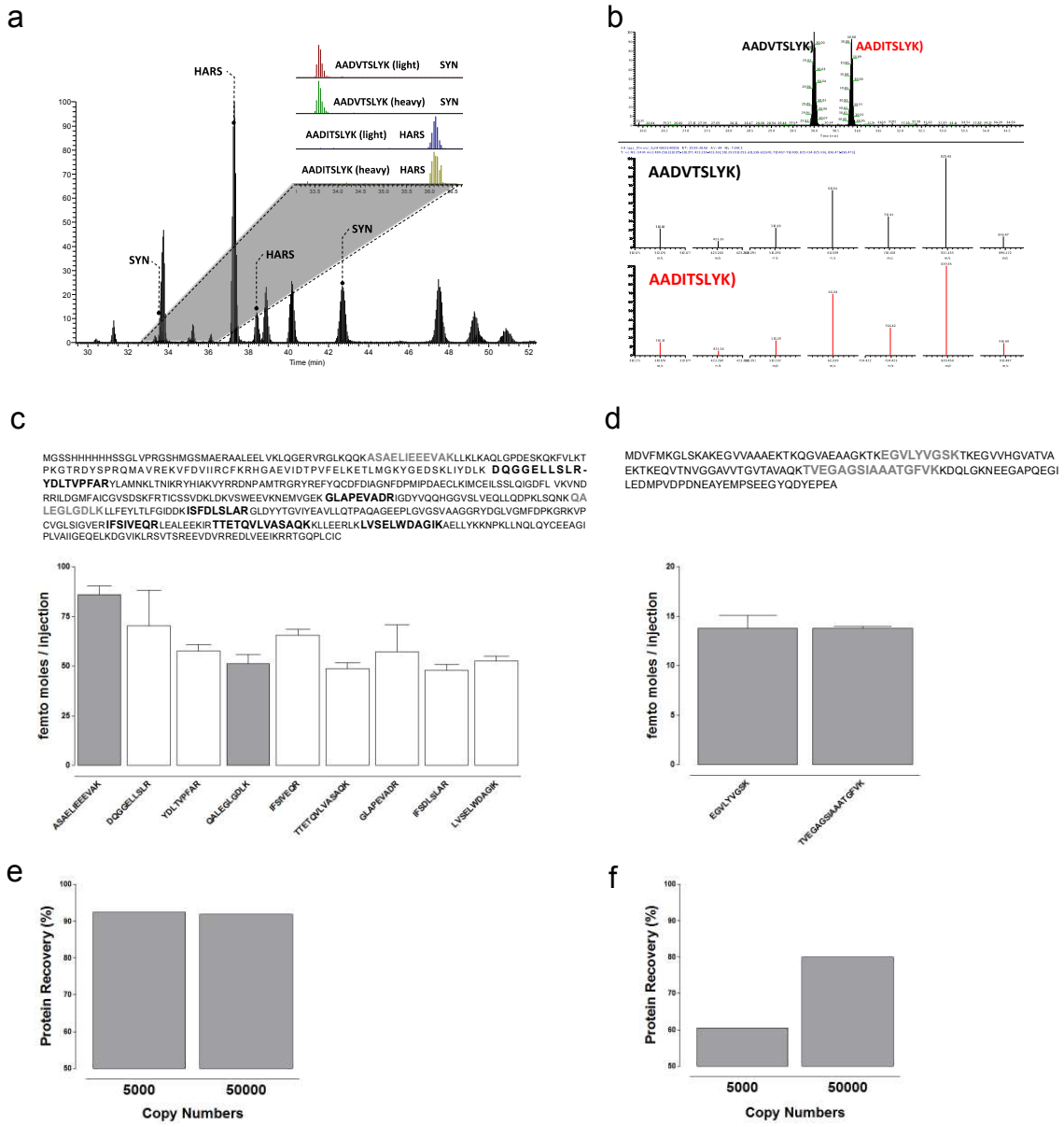
(**b**) Densitometric analysis was performed for both PPARγ (left) and RXRα (right). Values were normalized against PARP-1 and Day 0 was taken as reference. Values of the subsequent days are represented in terms of fold changes.

Supplementary Figure 9.



Supplementary Figure 9. Copy number assessment of RXRα in 3T3-L1 nuclear extracts by (semi-)quantitative Western blotting (WB).
(**a-b**) 3T3-L1 nuclear extracts of two independent biological replicates (Sample 1 and 2) were harvested at different time points during adipocyte differentiation (D0 to D6) and were Western blotted using RXRα-specific antibody. To generate a standard curve, different amounts of *in vitro*-expressed, accurately quantified (by SRM) recombinant RXRα were Western blotted using the same antibody (lanes 1-7 contain 20-0.31ng of recombinant RXRα), after which their respective band intensities were analyzed by spot densitometry. This revealed two regimes indicative of linearity up to 10 ng (blue line) after which saturation dampens the signal (red line). Samples from day one to four were therefore diluted 3-fold to operate in the linear regime after which the dilution factor was taken into account to calculate final numbers as shown in c). (**c**) Bar graph showing the average RXRα copy number at each differentiation time point based on Western blotting compared to that obtained by SRM. The number on top of the bar represents the fold change (FC) difference between the two assays. (**d**) Correlation between the copy number per cell calculated by SRM and Western blotting.

Supplementary Figure 10



Supplementary Figure 10. SRM assay validation using two accurately quantified proteins.

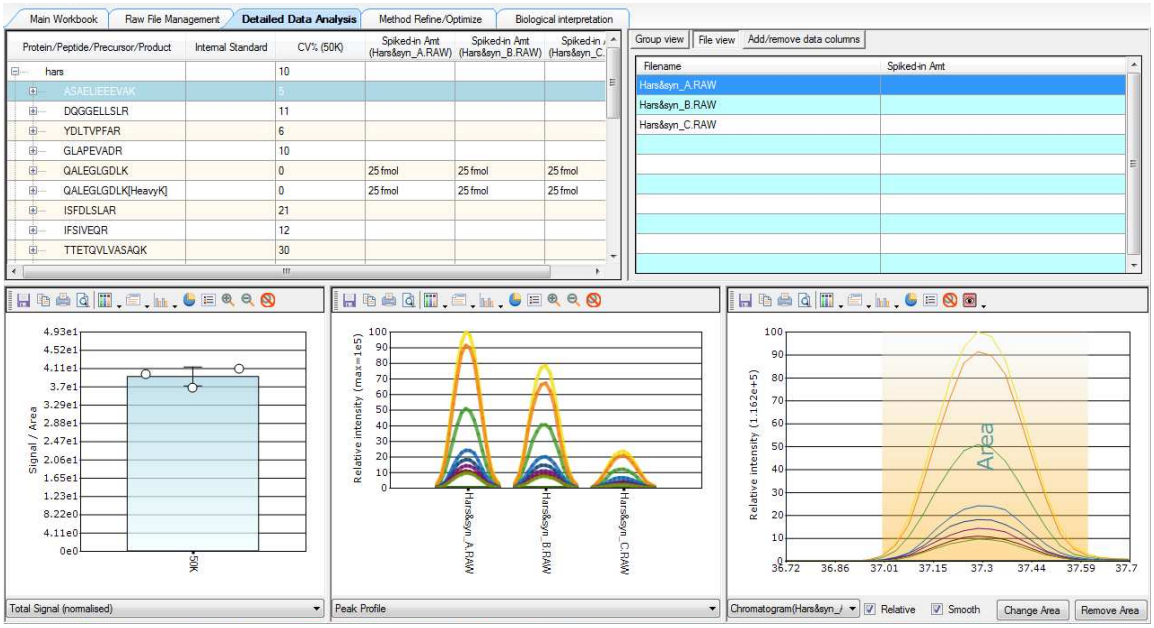
Recombinant human histidyl-tRNA synthetas (HARS) and human α-synuclein (SYN) were accurately quantified by UV absorbance and amino acid analysis (Supplementary Methods). Proteins were mixed with nuclear extract (D0), separated by SDS-PAGE and quantified using the SH-tag multiplex approach.

(**a**) LC-SRM spectrum of all HARS and SYN peptide fragments monitored by SRM. Proteotypic peptides for each protein are labeled as HARS and SYN respectively. Inset; total ion chromatogram (TIC) of light (spiked) and heavy labeled SH-tags for HARS and SYN, showing a clear distinction of both reference tags.

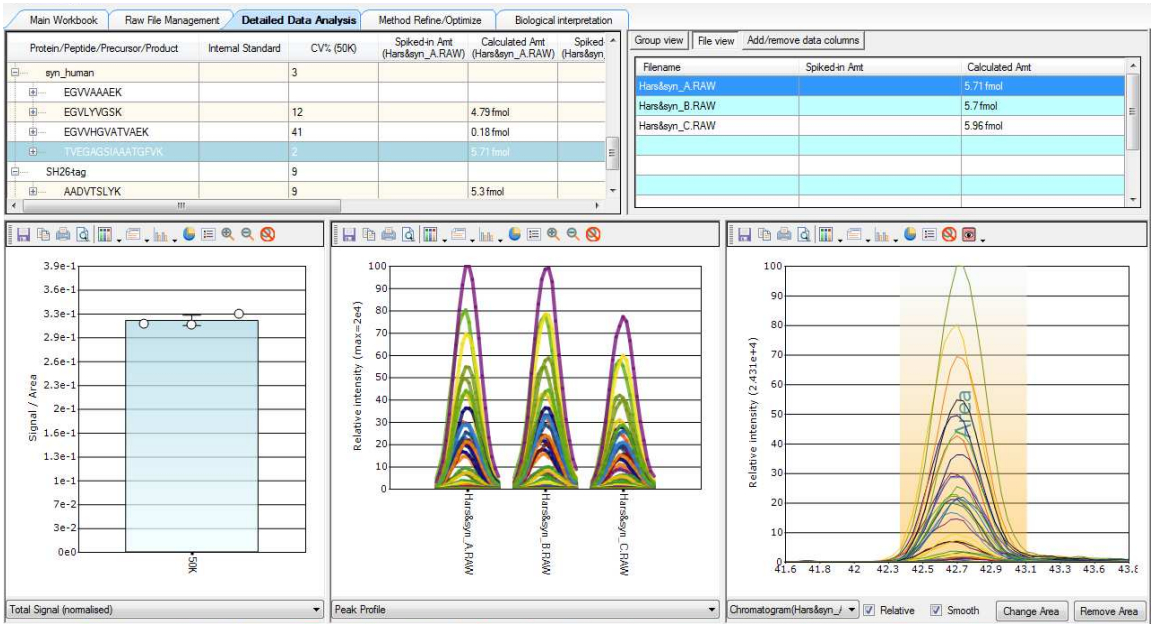
(**b**) The efficiency of SH-tag digestion and degree of recovery was assessed by monitoring the two SH-tags in a separate run. Recovery of both SH-tags was confirmed by identical signal intensities for parent- and fragment transition ions.
(**c**) Measured variation (mean ± SD, n=3) found with HARS peptides (peptides are corresponding to 50,000 copies per cell of HARS). A total of nine peptides were monitored for HARS protein of which two peptides are proteotypic (highlighted in grey).
(**d**) Measured variation of two proteotypic SYN peptides (mean ± SD, n=3; peptides are corresponding to 50,000 copies per cell of SYN). Endogenous (nuclear extract) HARS and SYN protein was assessed by monitoring mouse proteotypic peptides during each run.
(**e**) Percentage of the spiked-in 5,000 and 50,000 copies per cell of HARS which were recovered after in-gel digestion followed by SRM quantification.
(**f**) Percentage of the spiked-in 5,000 and 50,000 copies per cell of SYN which were recovered after in-gel digestion followed by SRM quantification.

Supplementary Figure 11.

a

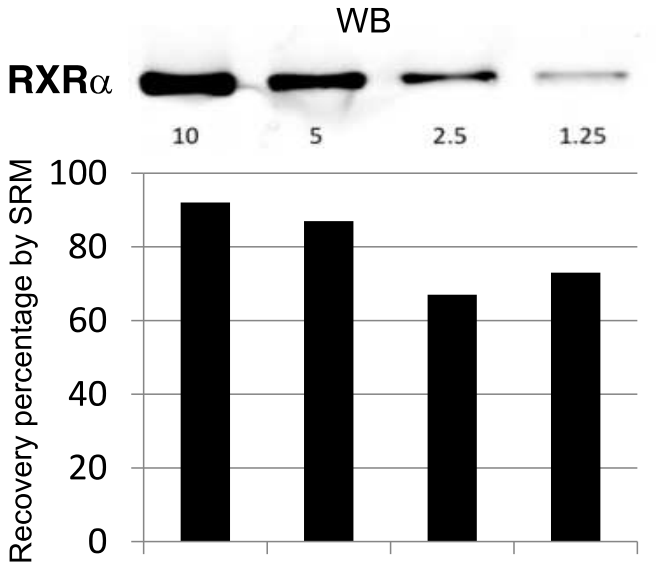


b



Supplementary Figure 11. Data evaluation using Pinpoint software.
The screenshots show an example of the data evaluation using Pinpoint software as well as the coefficient of variation (CV) found with a proteotypic peptide of HARS (**a**) and SYN (**b**) respectively. A CV of 5% and 2% was measured for a proteotypic peptide of HARS and SYN respectively.

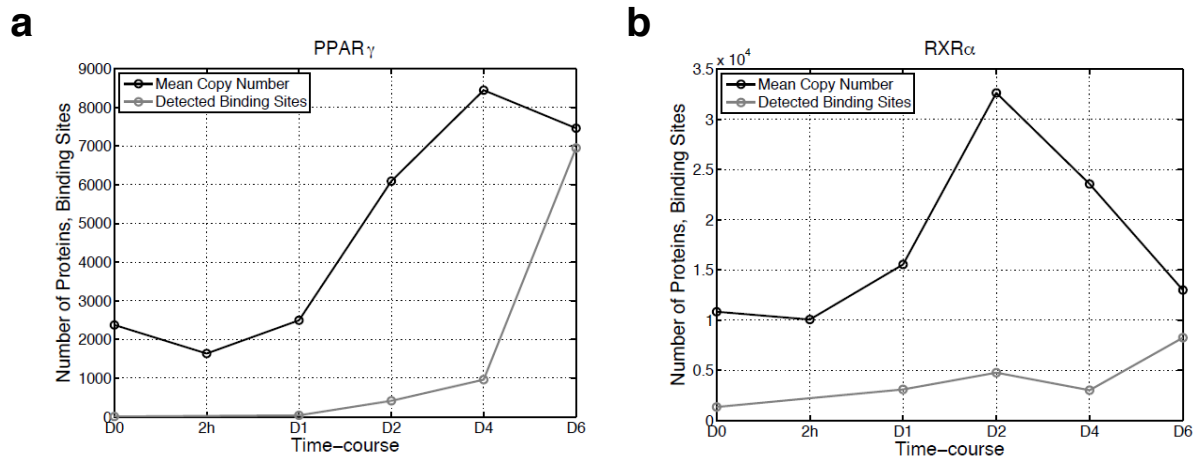
Supplementary Figure 12.



Supplementary Figure 12. **Assessment of the impact of gel extraction on protein quantification.**

In vitro-expressed, recombinant RXRα was quantified using SRM after which different amounts were resolved on an SDS-PAGE gel and Western blotted using an RXRα-specific antibody. A parallel series of corresponding bands were subsequently isolated and in-gel digested, after which the respective digests were subjected to SRM. The bar graph underneath shows the recovery percentage for each amount. These results indicate that for small amounts of TF (≤ 2.5 ng), protein loss after gel extraction, digestion etc. may amount to ~30%, whereas for larger amounts, the loss appears minimal (<~8%) as proteins likely function as carriers for themselves, consistent with the data derived from our HARS and SYN analyses.

Supplementary Figure 13.



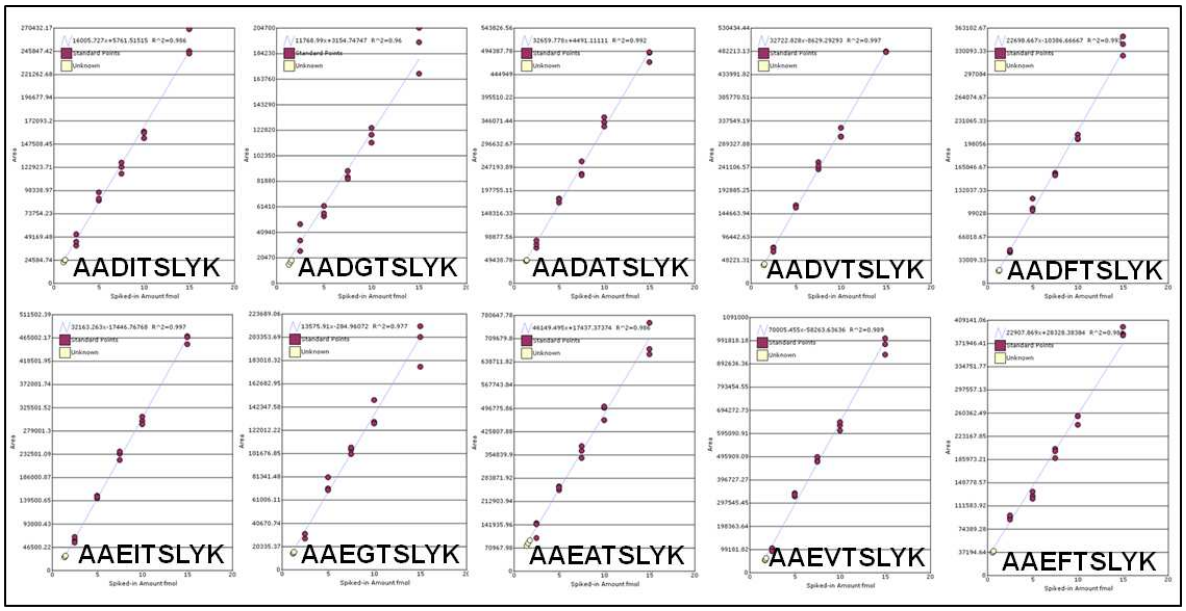
Supplementary Figure 13. PPARγ and RXRα protein copy number versus the number of detected binding sites.

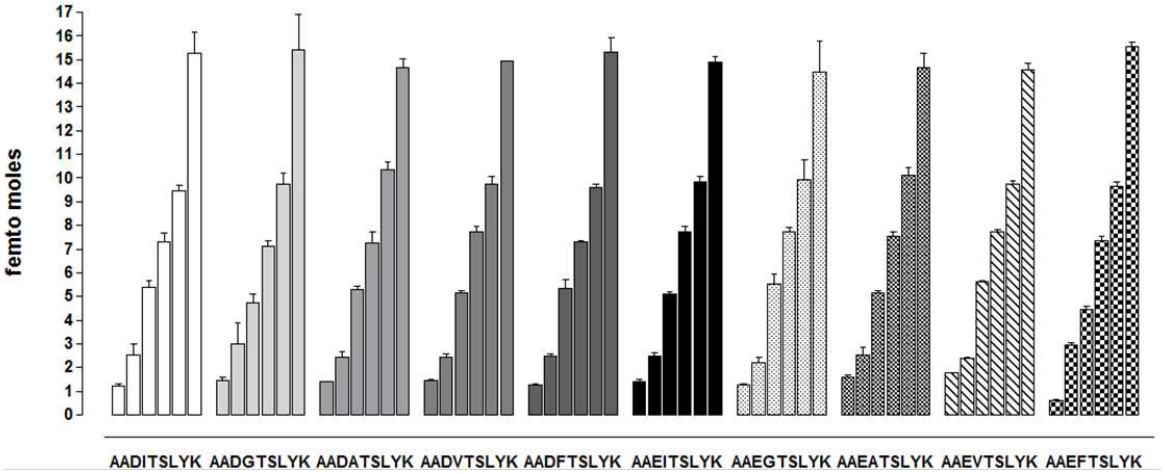
(**a**) Graph showing the PPARγ protein copy number profile during terminal adipogenesis as detected by SRMs in relation to its binding site profile. A sharp increase in detected binding sites can be observed after Day 4, despite a saturation in PPARγ protein copy numbers.

(**b**) Graph showing the RXRα protein copy number profile during terminal adipogenesis as detected by SRMs in relation to its binding site profile. An increase in detected binding sites can be observed after Day 4, despite a decrease in RXRα protein copy numbers.

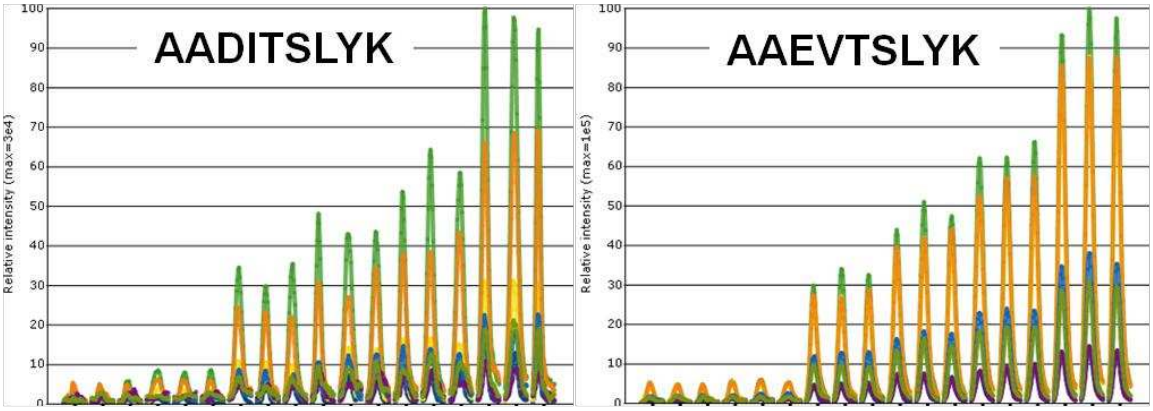
Supplementary Figure 14.

a





b



no.	Tag sequence	Parent Mass m/z $[M+2H]^{2+}$	CV of tags (1.25 - 15 fmole) 1.25	2.5	5.0	7.5	10	15
1	AADITSLYK	491.26	5%	11%	4%	4%	2%	5%
2	AADGTSLYK	463.24	8%	29%	6%	3%	4%	8%
3	AADATSLYK	470.25	2%	9%	2%	6%	2%	2%
4	AADVTSLYK	484.25	3%	6%	1%	3%	3%	1%
5	AADFTSLYK	508.25	4%	3%	7%	1%	1%	3%
6	AAEITSLYK	498.27	5%	5%	2%	3%	2%	2%
7	AAEGTSLYK	470.24	4%	7%	7%	2%	7%	8%
8	AAEATSLYK	477.25	8%	17%	2%	4%	4%	6%
9	AAEVTSLYK	491.26	7%	7%	2%	2%	2%	3%
10	AAEFTSLYK	515.26	3%	2%	4%	3%	3%	2%

Supplementary Figure 14. Establishment of standard curves for all ten SH-Quant tags used for multiplex SRM analysis.

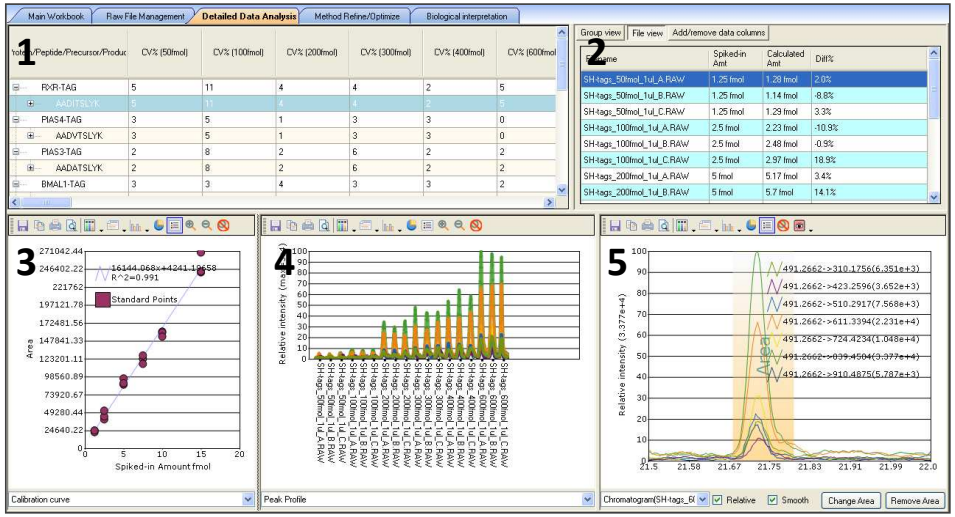
(**a**) All ten SH-Quant tags were spiked into nuclear extracts and analyzed in triplicates. Standard curves were established covering an on-column tag concentration range from 2.5 fmol to 15.0 fmol in a complex matrix. Top panels: linear regression analysis (R^2 = 0.960 – 0.997) for all ten SH-Quant tags. Bottom panel: graphical representation of the variation found within all technical replicates (2.5 - 15 fmol) as well as the calculated value of an unknown treated sample (positive control) spiked into nuclear extract at 1.25 fmol.

(**b**) Top panel: example of the calculated peak areas of two isobaric tags (tags 1 and 9), with slightly modified sequences, analyzed in three technical replicates at a concentration range of 1.25 – 15 fmol. Bottom panel: the table provides a summary of the calculated CV for all ten SH-Quant tags spiked at different concentrations (1.25 – 15 fmoles) into nuclear extracts. All samples were analyzed in triplicates.

Supplementary Figure 15.

a

AADITSLYK

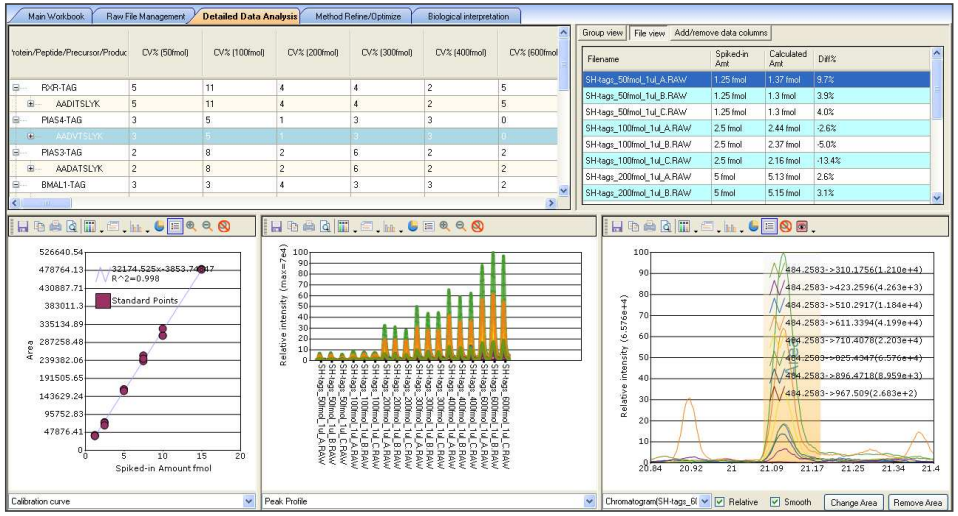


b

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.28 fmol	2.3%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.14 fmol	-8.5%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.3 fmol	3.7%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.23 fmol	-10.9%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.48 fmol	-0.9%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.97 fmol	18.8%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.16 fmol	3.2%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.69 fmol	13.9%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.28 fmol	5.7%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.34 fmol	-2.2%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	6.91 fmol	-7.9%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.63 fmol	1.7%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.65 fmol	-3.5%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.56 fmol	-4.4%
SH-tags_400fmol_1ul_C.RAW	10 fmol	9.22 fmol	-7.8%
SH-tags_600fmol_1ul_A.RAW	15 fmol	14.88 fmol	-0.8%
SH-tags_600fmol_1ul_B.RAW	15 fmol	14.73 fmol	-1.8%
SH-tags_600fmol_1ul_C.RAW	15 fmol	16.31 fmol	8.7%

c

AADVTSLYK

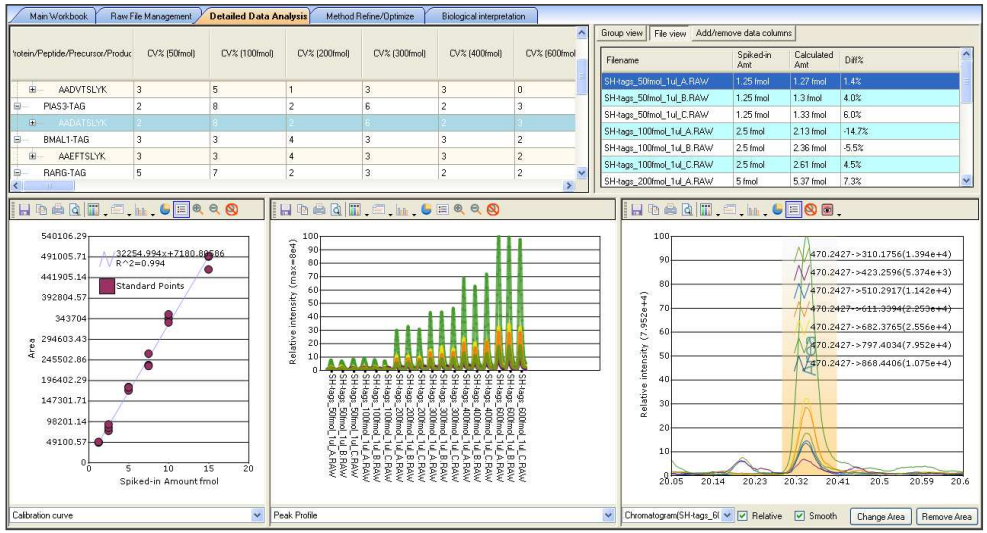


d

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.37 fmol	9.7%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.3 fmol	3.9%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.3 fmol	4.0%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.44 fmol	-2.6%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.37 fmol	-5.0%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.16 fmol	-13.4%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.13 fmol	2.6%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.15 fmol	3.1%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5 fmol	0.1%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.93 fmol	5.8%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.47 fmol	-0.4%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.67 fmol	2.3%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.58 fmol	-4.2%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.57 fmol	-4.3%
SH-tags_400fmol_1ul_C.RAW	10 fmol	10.15 fmol	1.5%
SH-tags_600fmol_1ul_A.RAW	15 fmol	15.07 fmol	0.5%
SH-tags_600fmol_1ul_B.RAW	15 fmol	15.05 fmol	0.3%
SH-tags_600fmol_1ul_C.RAW	15 fmol	15.02 fmol	0.1%

AADATSLYK

e

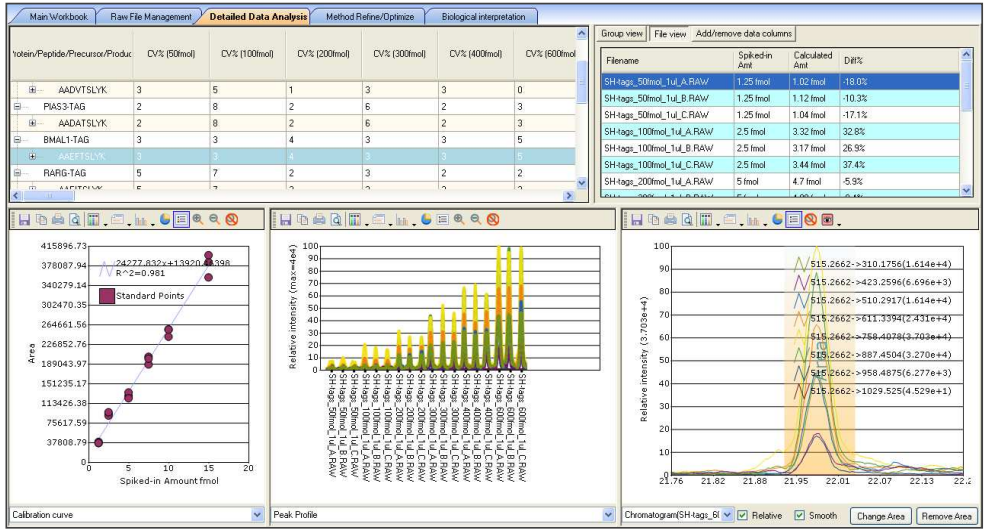


f

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.27 fmol	1.4%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.3 fmol	4.0%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.33 fmol	6.0%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.13 fmol	-14.7%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.36 fmol	-5.5%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.61 fmol	4.5%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.37 fmol	7.3%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.1 fmol	2.1%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.34 fmol	6.8%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	6.99 fmol	-6.8%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.82 fmol	4.2%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	6.91 fmol	-7.9%
SH-tags_400fmol_1ul_A.RAW	10 fmol	10.4 fmol	4.0%
SH-tags_400fmol_1ul_B.RAW	10 fmol	10.12 fmol	1.2%
SH-tags_400fmol_1ul_C.RAW	10 fmol	10.72 fmol	7.2%
SH-tags_600fmol_1ul_A.RAW	15 fmol	14.02 fmol	-6.5%
SH-tags_600fmol_1ul_B.RAW	15 fmol	14.97 fmol	-0.2%
SH-tags_600fmol_1ul_C.RAW	15 fmol	15 fmol	0.0%

AAEFTSLYK

g

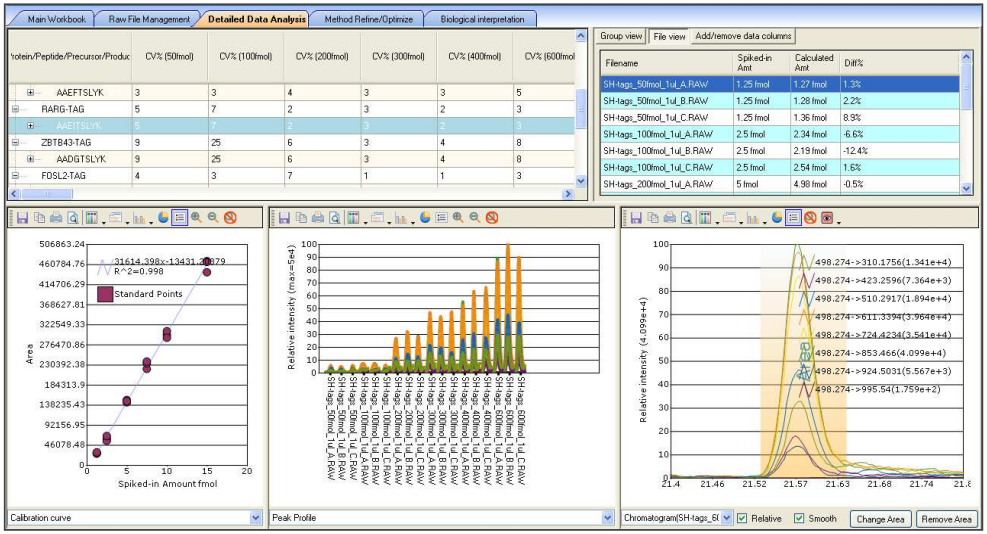


h

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.02 fmol	-18.0%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.12 fmol	-10.3%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.04 fmol	-17.1%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	3.32 fmol	32.8%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	3.17 fmol	26.9%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	3.44 fmol	37.4%
SH-tags_200fmol_1ul_A.RAW	5 fmol	4.7 fmol	-5.9%
SH-tags_200fmol_1ul_B.RAW	5 fmol	4.98 fmol	-0.4%
SH-tags_200fmol_1ul_C.RAW	5 fmol	4.52 fmol	-9.6%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.16 fmol	-4.5%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.76 fmol	3.5%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.61 fmol	1.5%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.89 fmol	-1.1%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.89 fmol	-1.1%
SH-tags_400fmol_1ul_C.RAW	10 fmol	9.33 fmol	-6.7%
SH-tags_600fmol_1ul_A.RAW	15 fmol	13.96 fmol	-7.0%
SH-tags_600fmol_1ul_B.RAW	15 fmol	15.14 fmol	0.9%
SH-tags_600fmol_1ul_C.RAW	15 fmol	15.69 fmol	4.6%

i

AAEITSLYK

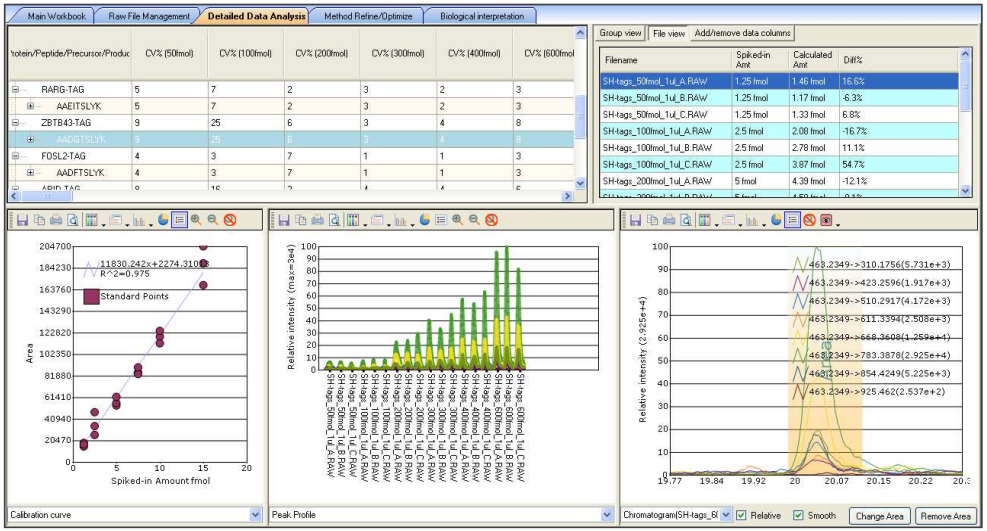


j

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.27 fmol	1.3%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.28 fmol	2.2%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.36 fmol	8.9%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.34 fmol	-6.6%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.19 fmol	-12.4%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.54 fmol	1.6%
SH-tags_200fmol_1ul_A.RAW	5 fmol	4.98 fmol	-0.5%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.15 fmol	3.0%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.07 fmol	1.5%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.4 fmol	-1.3%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.8 fmol	4.0%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.94 fmol	5.8%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.87 fmol	-1.3%
SH-tags_400fmol_1ul_B.RAW	10 fmol	10.14 fmol	1.4%
SH-tags_400fmol_1ul_C.RAW	10 fmol	9.66 fmol	-3.4%
SH-tags_600fmol_1ul_A.RAW	15 fmol	14.41 fmol	-4.0%
SH-tags_600fmol_1ul_B.RAW	15 fmol	15.22 fmol	1.4%
SH-tags_600fmol_1ul_C.RAW	15 fmol	15.16 fmol	1.0%

k

AADGTSLYK

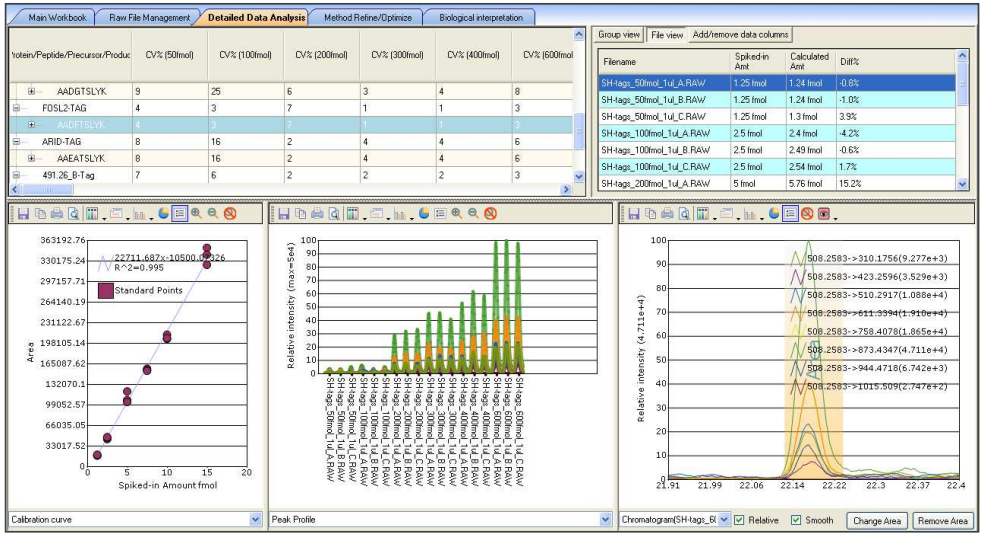


l

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.46 fmol	16.6%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.17 fmol	-6.3%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.33 fmol	6.8%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.08 fmol	-16.7%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.78 fmol	11.1%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	3.87 fmol	54.7%
SH-tags_200fmol_1ul_A.RAW	5 fmol	4.39 fmol	-12.1%
SH-tags_200fmol_1ul_B.RAW	5 fmol	4.59 fmol	-8.1%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.08 fmol	1.5%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7 fmol	-6.6%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	6.87 fmol	-8.4%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.4 fmol	-1.3%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.81 fmol	-1.9%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.29 fmol	-7.1%
SH-tags_400fmol_1ul_C.RAW	10 fmol	10.26 fmol	2.6%
SH-tags_600fmol_1ul_A.RAW	15 fmol	15.59 fmol	3.9%
SH-tags_600fmol_1ul_B.RAW	15 fmol	16.91 fmol	12.7%
SH-tags_600fmol_1ul_C.RAW	15 fmol	13.87 fmol	-7.6%

AADFTSLYK

m

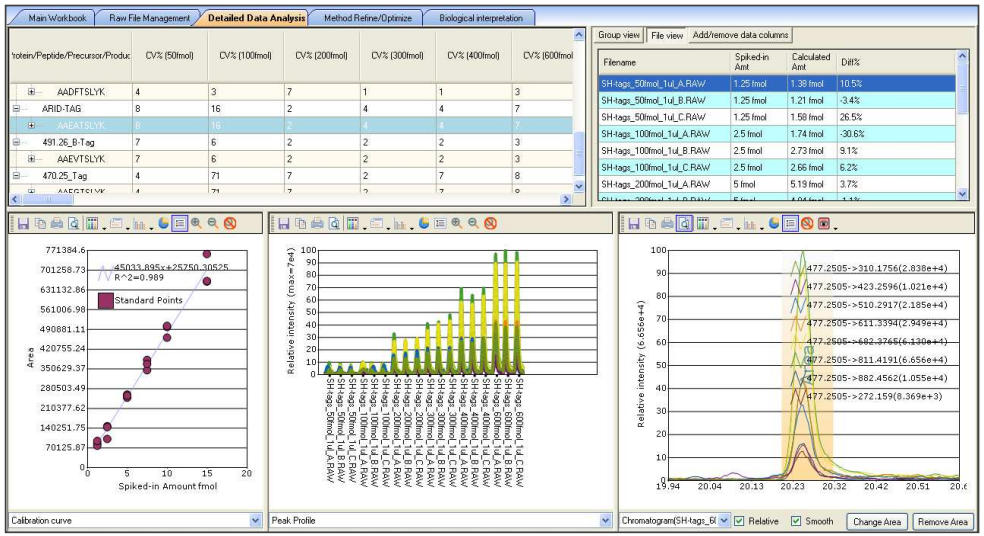


n

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.24 fmol	-0.8%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.24 fmol	-1.0%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.3 fmol	3.9%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.4 fmol	-4.2%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.49 fmol	-0.6%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.54 fmol	1.7%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.76 fmol	15.2%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.15 fmol	3.0%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.01 fmol	0.3%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.37 fmol	-1.8%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.26 fmol	-3.1%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.22 fmol	-3.8%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.77 fmol	-2.3%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.47 fmol	-5.3%
SH-tags_400fmol_1ul_C.RAW	10 fmol	9.53 fmol	-4.7%
SH-tags_600fmol_1ul_A.RAW	15 fmol	14.7 fmol	-2.0%
SH-tags_600fmol_1ul_B.RAW	15 fmol	15.9 fmol	6.0%
SH-tags_600fmol_1ul_C.RAW	15 fmol	15.42 fmol	2.8%

AAEATSLYK

o

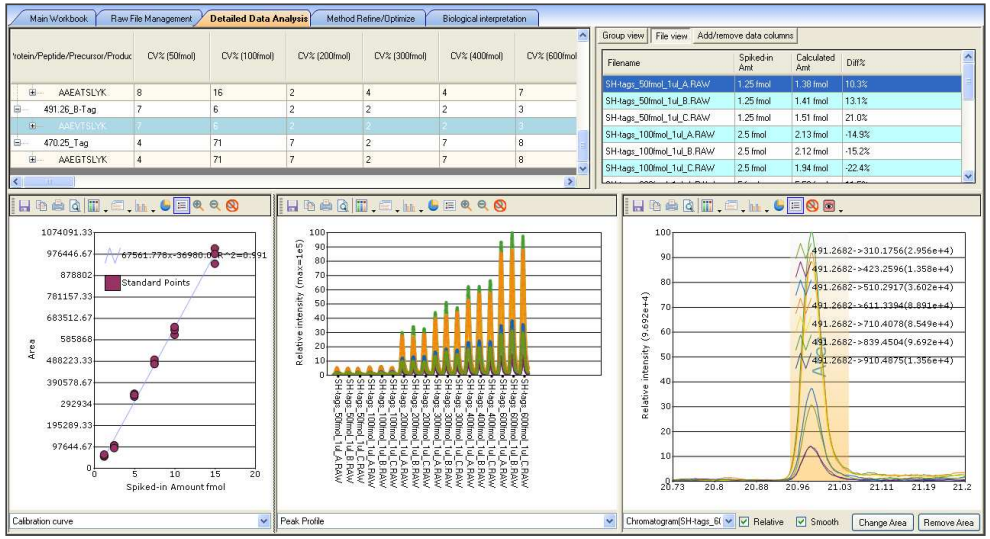


p

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.38 fmol	10.5%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.21 fmol	-3.4%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.58 fmol	26.5%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	1.74 fmol	-30.6%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.73 fmol	9.1%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	2.66 fmol	6.2%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.19 fmol	3.7%
SH-tags_200fmol_1ul_B.RAW	5 fmol	4.94 fmol	-1.1%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.11 fmol	2.1%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.08 fmol	-5.5%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.88 fmol	5.1%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.57 fmol	0.9%
SH-tags_400fmol_1ul_A.RAW	10 fmol	10.44 fmol	4.4%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.61 fmol	-3.9%
SH-tags_400fmol_1ul_C.RAW	10 fmol	10.53 fmol	5.3%
SH-tags_600fmol_1ul_A.RAW	15 fmol	13.98 fmol	-6.8%
SH-tags_600fmol_1ul_B.RAW	15 fmol	16.12 fmol	7.5%
SH-tags_600fmol_1ul_C.RAW	15 fmol	14.02 fmol	-6.6%

AAEVTSLYK

q

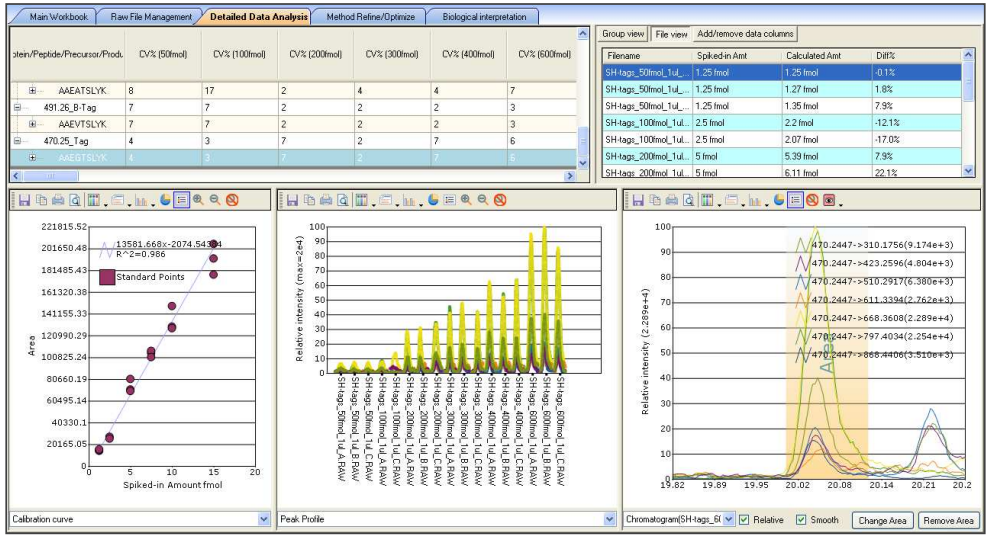


r

SH-tags_50fmol_1ul_A.RAW	1.25 fmol	1.38 fmol	10.3%
SH-tags_50fmol_1ul_B.RAW	1.25 fmol	1.41 fmol	13.1%
SH-tags_50fmol_1ul_C.RAW	1.25 fmol	1.51 fmol	21.0%
SH-tags_100fmol_1ul_A.RAW	2.5 fmol	2.13 fmol	-14.9%
SH-tags_100fmol_1ul_B.RAW	2.5 fmol	2.12 fmol	-15.2%
SH-tags_100fmol_1ul_C.RAW	2.5 fmol	1.94 fmol	-22.4%
SH-tags_200fmol_1ul_A.RAW	5 fmol	5.58 fmol	11.5%
SH-tags_200fmol_1ul_B.RAW	5 fmol	5.38 fmol	7.6%
SH-tags_200fmol_1ul_C.RAW	5 fmol	5.49 fmol	9.9%
SH-tags_300fmol_1ul_A.RAW	7.5 fmol	7.6 fmol	1.4%
SH-tags_300fmol_1ul_B.RAW	7.5 fmol	7.85 fmol	4.7%
SH-tags_300fmol_1ul_C.RAW	7.5 fmol	7.55 fmol	0.7%
SH-tags_400fmol_1ul_A.RAW	10 fmol	9.51 fmol	-4.9%
SH-tags_400fmol_1ul_B.RAW	10 fmol	9.83 fmol	-1.7%
SH-tags_400fmol_1ul_C.RAW	10 fmol	10.03 fmol	0.3%
SH-tags_600fmol_1ul_A.RAW	15 fmol	14.25 fmol	-5.0%
SH-tags_600fmol_1ul_B.RAW	15 fmol	15.28 fmol	1.8%
SH-tags_600fmol_1ul_C.RAW	15 fmol	14.91 fmol	-0.6%

AAEGTSLYK

s



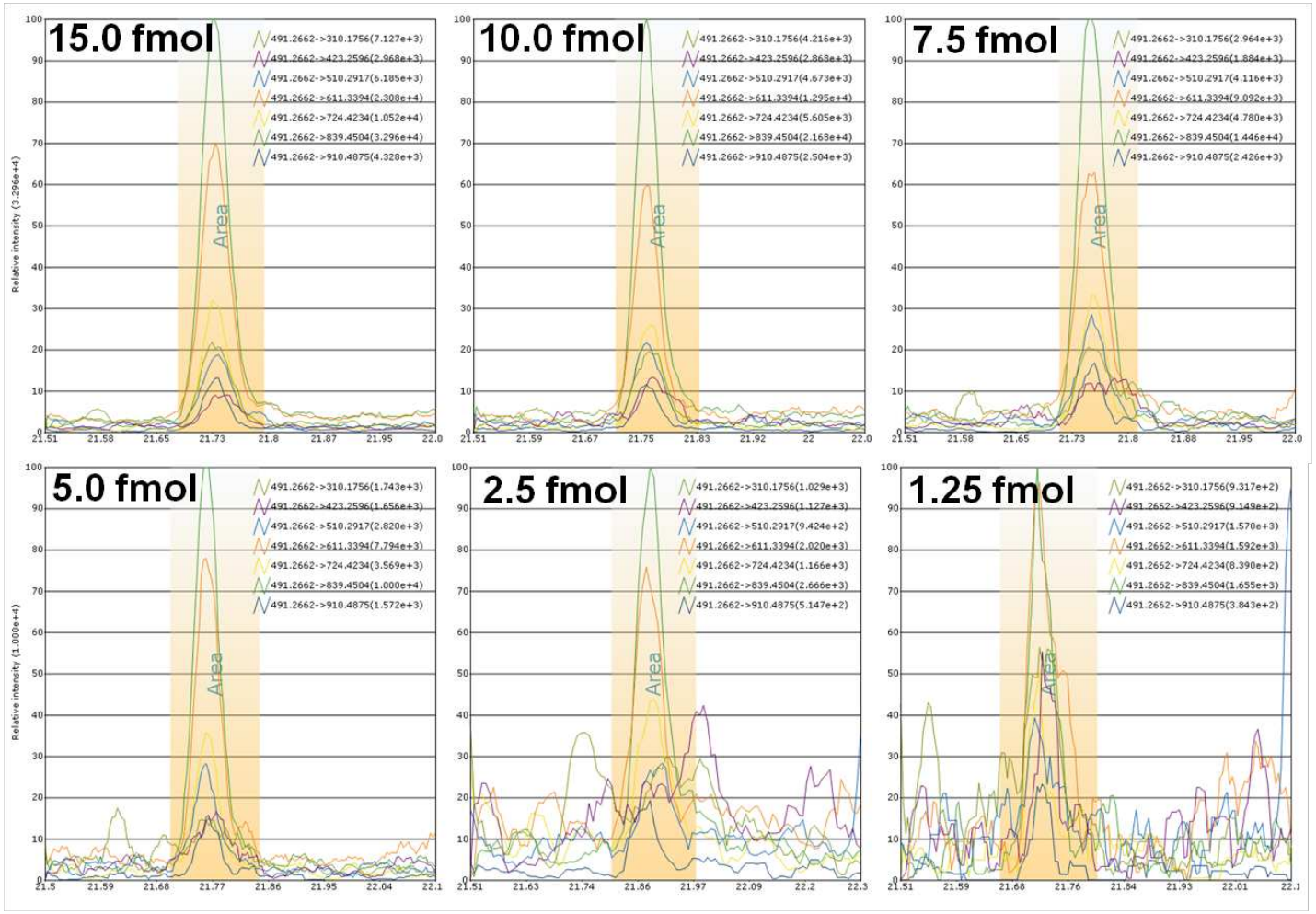
t

SH-tags_50fmol_1ul_...	1.25 fmol	1.25 fmol	-0.1%
SH-tags_50fmol_1ul_...	1.25 fmol	1.27 fmol	1.8%
SH-tags_50fmol_1ul_...	1.25 fmol	1.35 fmol	7.9%
SH-tags_100fmol_1ul...	2.5 fmol	2.2 fmol	-12.1%
SH-tags_100fmol_1ul...	2.5 fmol	2.07 fmol	-17.0%
SH-tags_200fmol_1ul...	5 fmol	5.39 fmol	7.9%
SH-tags_200fmol_1ul...	5 fmol	6.11 fmol	22.1%
SH-tags_200fmol_1ul...	5 fmol	5.29 fmol	5.8%
SH-tags_300fmol_1ul...	7.5 fmol	7.86 fmol	4.8%
SH-tags_300fmol_1ul...	7.5 fmol	8.01 fmol	6.8%
SH-tags_300fmol_1ul...	7.5 fmol	7.58 fmol	1.1%
SH-tags_400fmol_1ul...	10 fmol	9.64 fmol	-3.6%
SH-tags_400fmol_1ul...	10 fmol	9.54 fmol	-4.6%
SH-tags_400fmol_1ul...	10 fmol	11.03 fmol	10.3%
SH-tags_600fmol_1ul...	15 fmol	14.25 fmol	-5.0%
SH-tags_600fmol_1ul...	15 fmol	15.24 fmol	1.6%
SH-tags_600fmol_1ul...	15 fmol	13.16 fmol	-12.3%

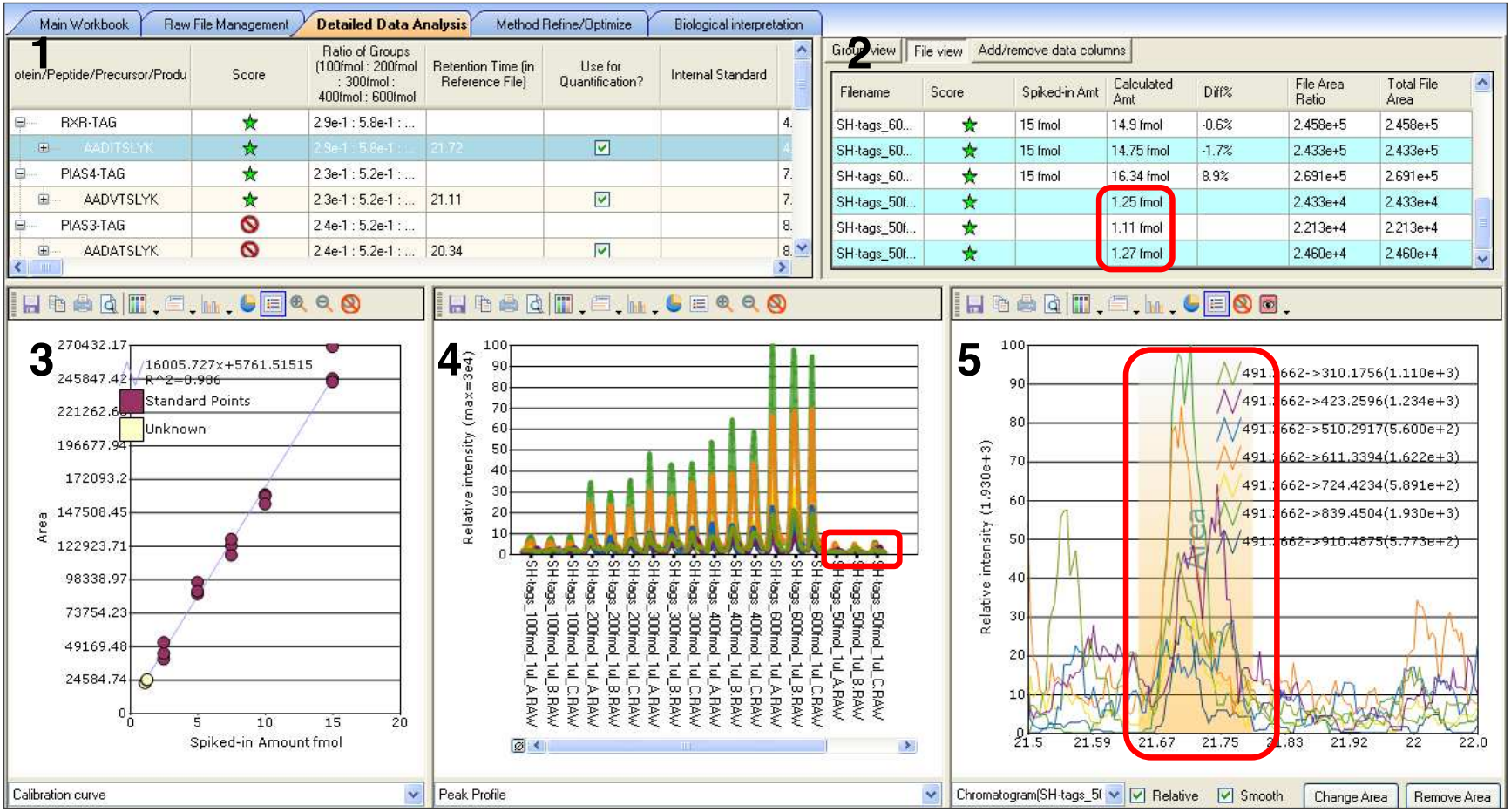
Supplementary Figure 15. Standard curves obtained with the ten SH-Quant tags in 3T3-L1 nuclear extract.
(**a,c,e,g,i,k,m,o,q,s**) Standard curves of all ten SH-Quant tags. The Pinpoint window provides a summary of the evaluation of the peptide-tags of interest. Subpanels 1 and 2 show an overview of the coefficients of variation per peptide-tag per time-point. Subpanel 3 shows the calculated standard curve regression line. Subpanel 4 shows the peak profiles of all technical replicates. Subpanel 5 shows the calculated peak area and chromatogram of peptide-tags transitions. Panel legend and numbering applies to all figures. (**b,d,f,h,j,l,n,p,r,t**) Calculated amounts of spiked samples.

Supplementary Figure 16.

a



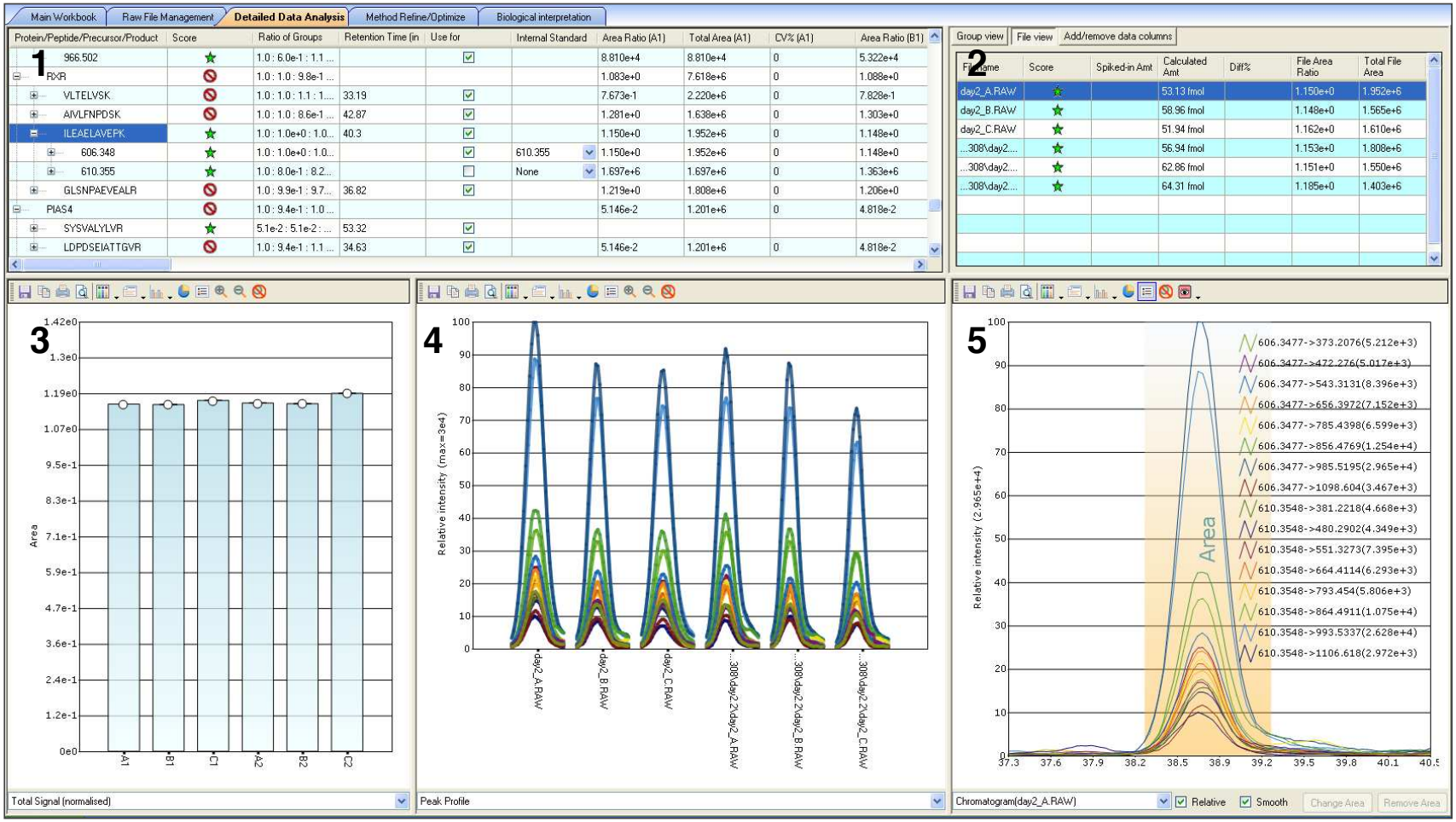
b


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
1
otein/Peptide/Precursor/Produ
Score
Ratio of Groups (100fmol : 200fmol : 300fmol : 400fmol : 600fmol
Retention Time (in Reference File)
Use for Quantification?
Internal Standard
RXR-TAG
AADITSLYK
PIAS4-TAG
AADVTSLYK
PIAS3-TAG
AADATSLYK
2.9e-1 : 5.8e-1 : ...
2.3e-1 : 5.2e-1 : ...
2.4e-1 : 5.2e-1 : ...
21.72
21.11
20.34
2
File view
Add/remove data columns
Filename
Score
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
SH-tags_60...
SH-tags_50f...
15 fmol
1.25 fmol
1.11 fmol
1.27 fmol
14.9 fmol
14.75 fmol
16.34 fmol
-0.6%
-1.7%
8.9%
2.458e+5
2.433e+5
2.691e+5
2.433e+4
2.213e+4
2.460e+4
3
16005.727x+5761.51515
R^2=0.986
Standard Points
Unknown
Area
Spiked-in Amount fmol
Calibration curve
4
Relative intensity (max=3e4)
Peak Profile
5
Relative intensity (1.930e+3)
662->310.1756(1.110e+3)
662->423.2596(1.234e+3)
662->510.2917(5.600e+2)
662->611.3394(1.622e+3)
662->724.4234(5.891e+2)
662->839.4504(1.930e+3)
662->910.4875(5.773e+2)
Chromatogram(SH-tags_5(
Relative
Smooth
Change Area
Remove Area

Supplementary Figure 16. Determination of the lower limit of quantitation (LLOQ) of SH-Quant tags in multiplexed samples.
(**a**) Example of the calculated peak area of a SH-Quant tag (AADITSLYK) spiked at concentrations of 15.0 fmol, 5.0 fmol, 2.50 fmol and 1.25 fmol (unknown treated sample). Little variation (≤ 0.15 min) in elution time was observed for all monitored concentrations.
(**b**) The Pinpoint window provides a summary of the evaluation of the peptide-tags of interest. Subpanel 1 presents an overview of the all peptides monitored. Subpanel 2 shows the calculated peptide-tag amounts. Subpanel 3 shows the calculated standard curve regression line. Subpanel 4 shows peak profiles of different spikes of all technical replicates. Subpanel 5 shows the calculated peak area and chromatogram of the transitions of a lower-end peptide-tag. The calculated concentrations of all data points within the calibration curve were found to lie in the acceptable range of ± 15% variation. The red boxes highlight the peak area and concentrations found for the “unknown” sample. The three technical replicates show little variation in concentrations (1.25 fmol, 1.11 fmol, 1.27 fmol, mean = 1.21 ± 0.09, variation ± 7.4%), which fulfills the acceptable range of variation as set by the FDA guidelines for the LLOQ (± 20%). A complete overview of all ten SH-Quant tags with the calculated concentrations can be found in Supplementary Figure 10.

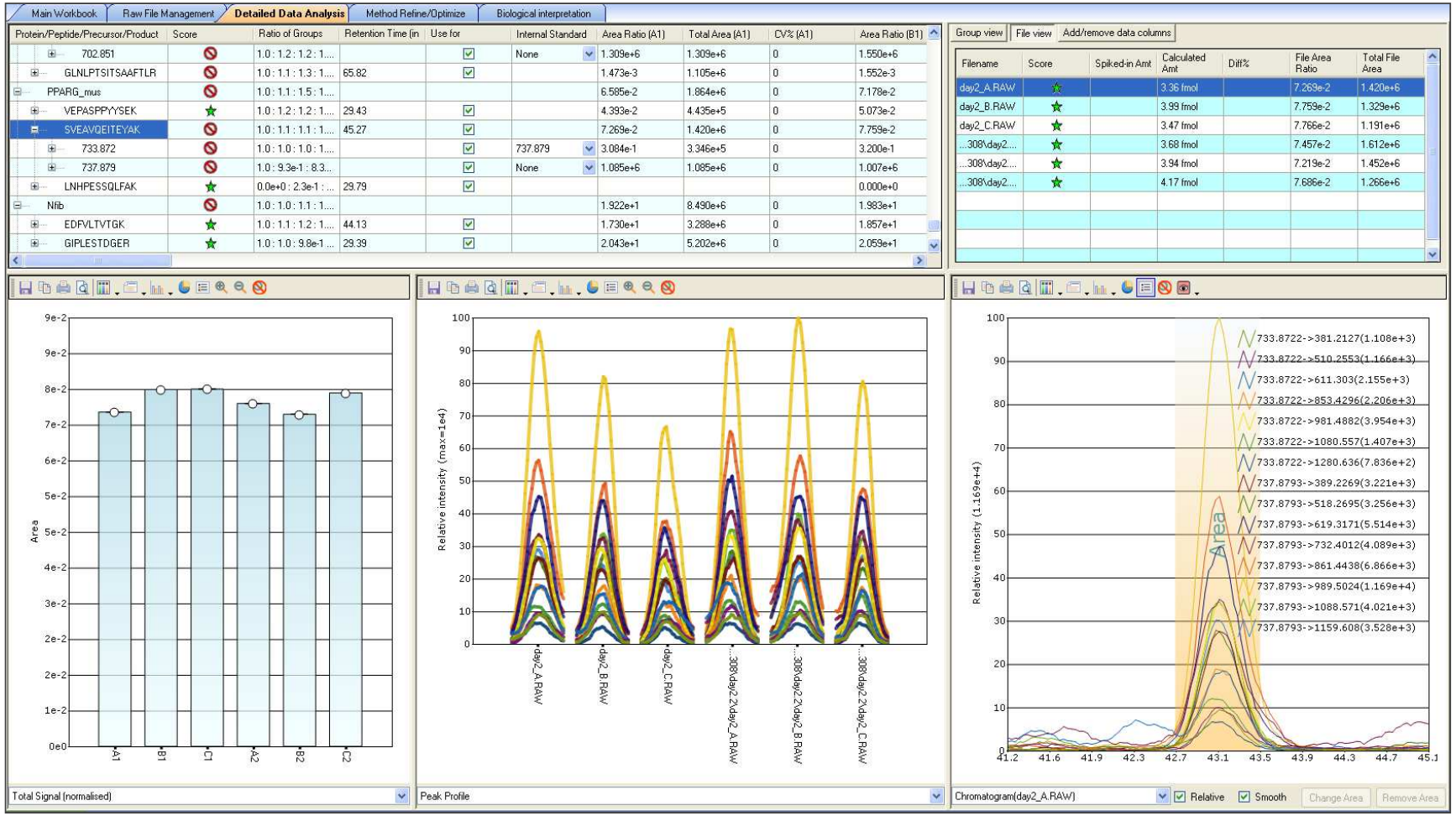
Supplementary Figure 17.

a **RXRα**



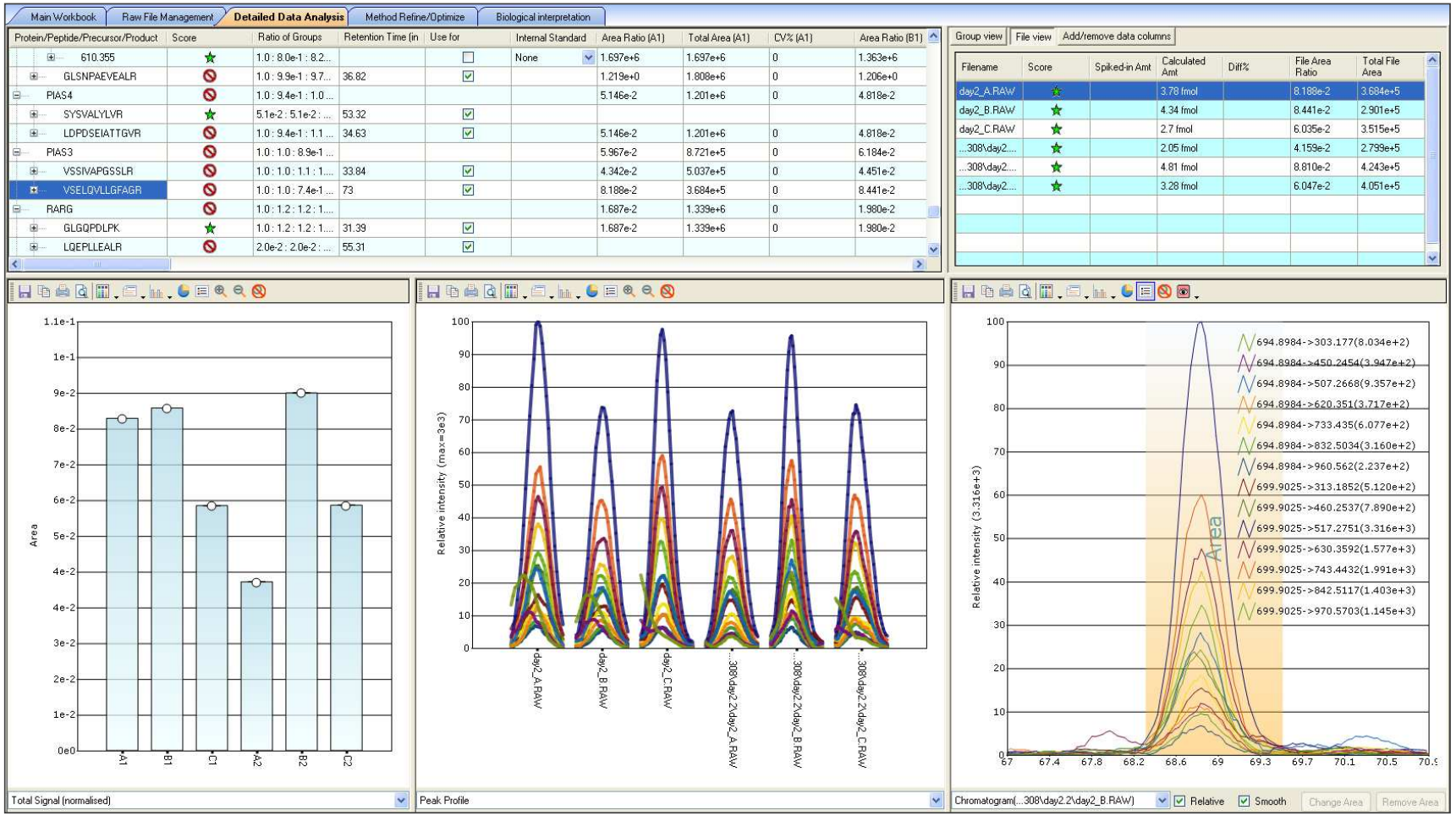
b

PPARγ


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
702.851
GLNLPTSITSAAFTLR
PPARG_mus
VEPASPPYYSEK
SVEAVQEITEYAK
733.872
737.879
LNHPESSQLFAK
Nfib
EDFVLTVTGK
GIPLESTDGER
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
Area
Relative intensity (max=1e4)
Relative intensity (1.169e+4)
733.8722->381.2127(1.108e+3)
733.8722->510.2553(1.166e+3)
733.8722->611.303(2.155e+3)
733.8722->853.4296(2.206e+3)
733.8722->981.4882(3.954e+3)
733.8722->1080.557(1.407e+3)
733.8722->1280.636(7.836e+2)
737.8793->389.2269(3.221e+3)
737.8793->518.2695(3.256e+3)
737.8793->619.3171(5.514e+3)
737.8793->732.4012(4.089e+3)
737.8793->861.4438(6.866e+3)
737.8793->989.5024(1.169e+4)
737.8793->1088.571(4.021e+3)
737.8793->1159.608(3.528e+3)
Total Signal (normalised)
Peak Profile
Chromatogram(day2_A.RAW)
Relative
Smooth
Change Area
Remove Area

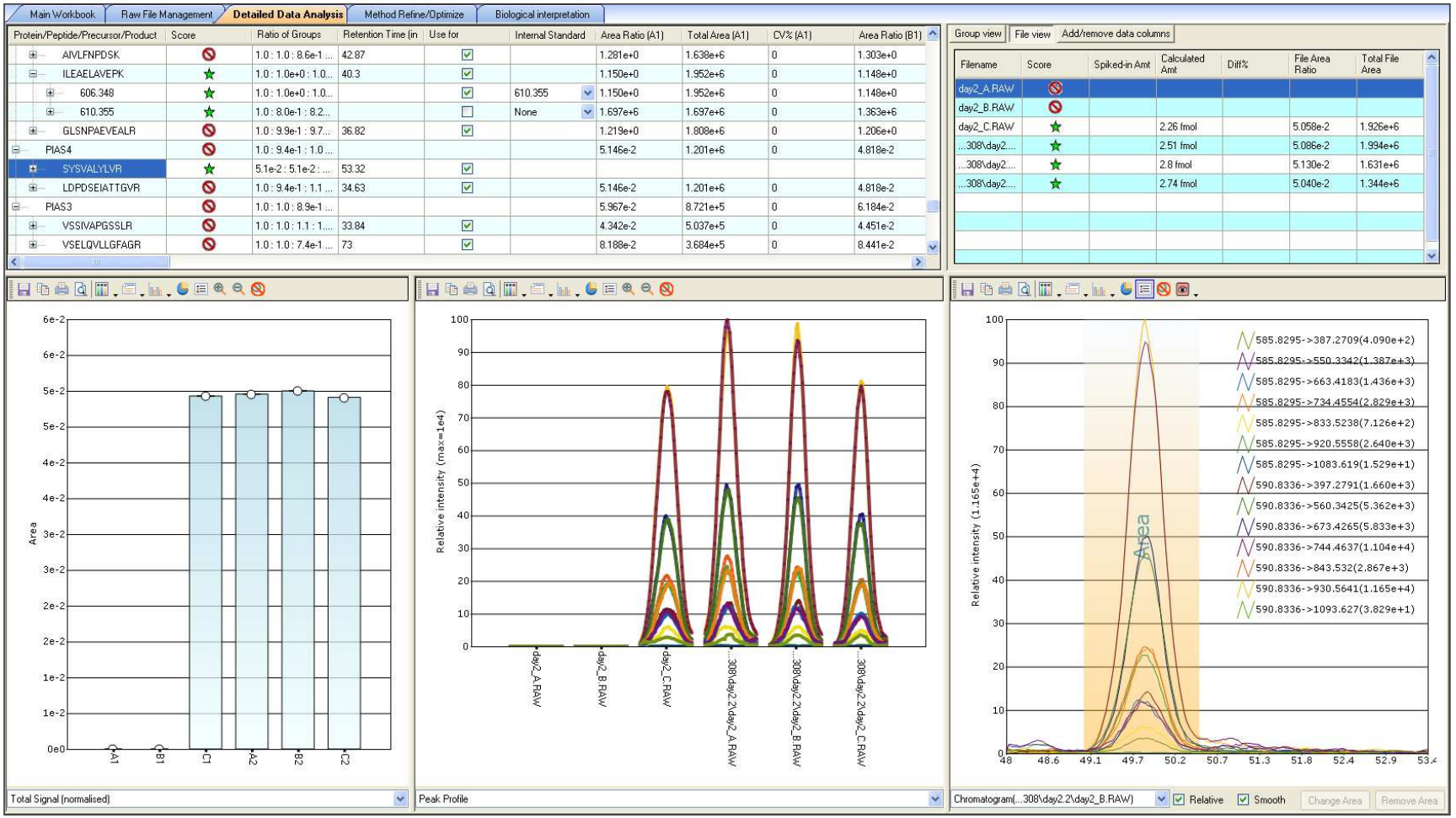
c

PIAS3


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
610.355
GLSNPAEVEALR
PIAS4
SYSVALYLVR
LDPDSEIATTGVR
PIAS3
VSSIVAPGSSLR
VSELQVLLGFAGR
RARG
GLGQPDLPK
LQEPLLEALR
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
Area
Relative intensity (max=3e3)
Relative intensity (3.316e+3)
694.8984->303.177(8.034e+2)
694.8984->450.2454(3.947e+2)
694.8984->507.2668(9.357e+2)
694.8984->620.351(3.717e+2)
694.8984->733.435(6.077e+2)
694.8984->832.5034(3.160e+2)
694.8984->960.562(2.237e+2)
699.9025->313.1852(5.120e+2)
699.9025->460.2537(7.890e+2)
699.9025->517.2751(3.316e+3)
699.9025->630.3592(1.577e+3)
699.9025->743.4432(1.991e+3)
699.9025->842.5117(1.403e+3)
699.9025->970.5703(1.145e+3)
Total Signal (normalised)
Peak Profile
Relative
Smooth
Change Area
Remove Area

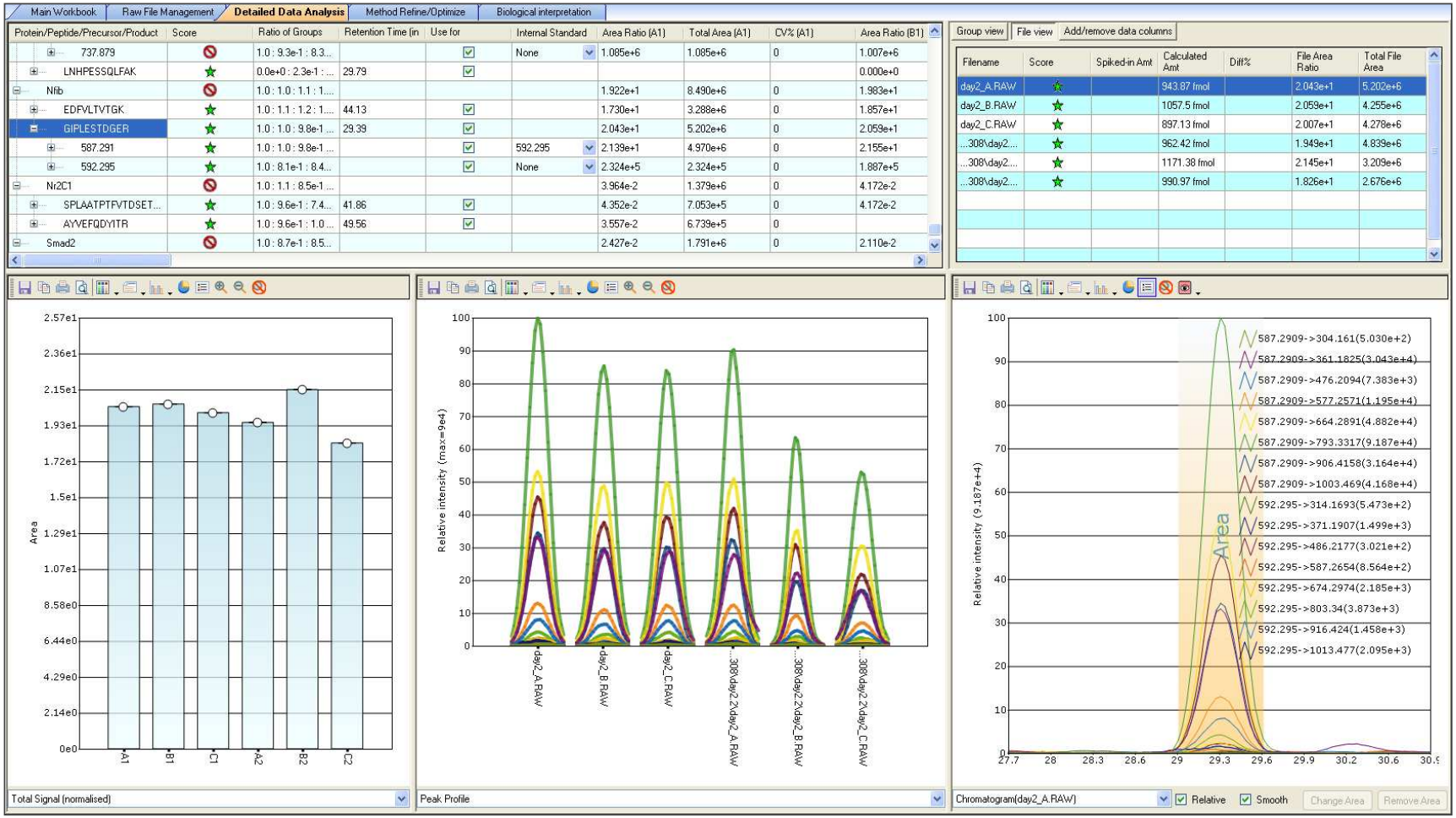
d

PIAS4


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
AIVLFNPDSK
ILEAELAVEPK
606.348
610.355
GLSNPAEVEALR
PIAS4
SYSVALYLVR
LDPDSEIATTGVR
PIAS3
VSSIVAPGSSLR
VSELQVLLGFAGR
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
2.26 fmol
2.51 fmol
2.8 fmol
2.74 fmol
Area
Relative intensity (max=1e4)
Relative intensity (1.165e+4)
585.8295->387.2709(4.090e+2)
585.8295->550.3342(1.387e+3)
585.8295->663.4183(1.436e+3)
585.8295->734.4554(2.829e+3)
585.8295->833.5238(7.126e+2)
585.8295->920.5558(2.640e+3)
585.8295->1083.619(1.529e+1)
590.8336->397.2791(1.660e+3)
590.8336->560.3425(5.362e+3)
590.8336->673.4265(5.833e+3)
590.8336->744.4637(1.104e+4)
590.8336->843.532(2.867e+3)
590.8336->930.5641(1.165e+4)
590.8336->1093.627(3.829e+1)
Total Signal (normalised)
Peak Profile
Relative
Smooth
Change Area
Remove Area

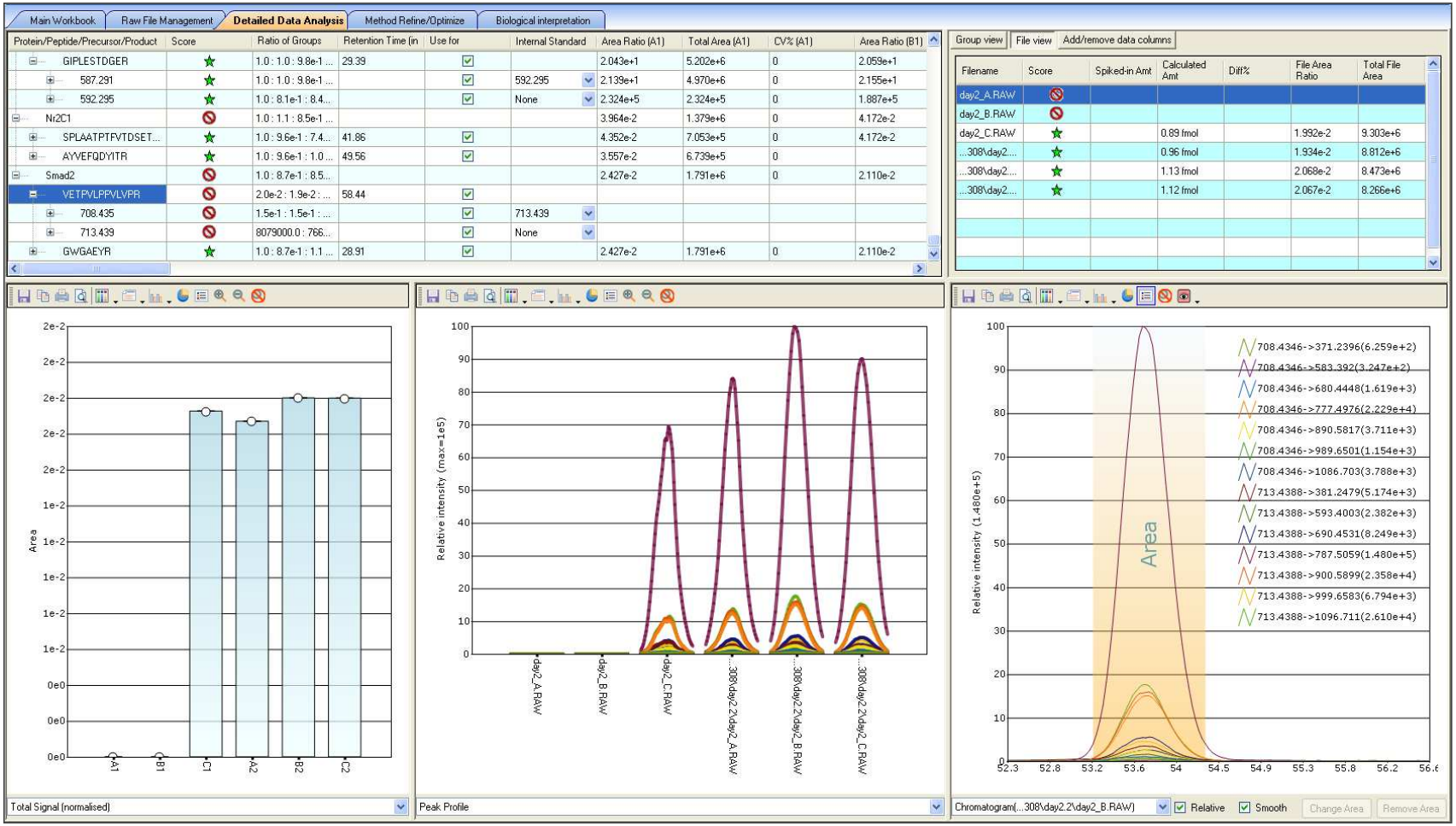
e

NFIB

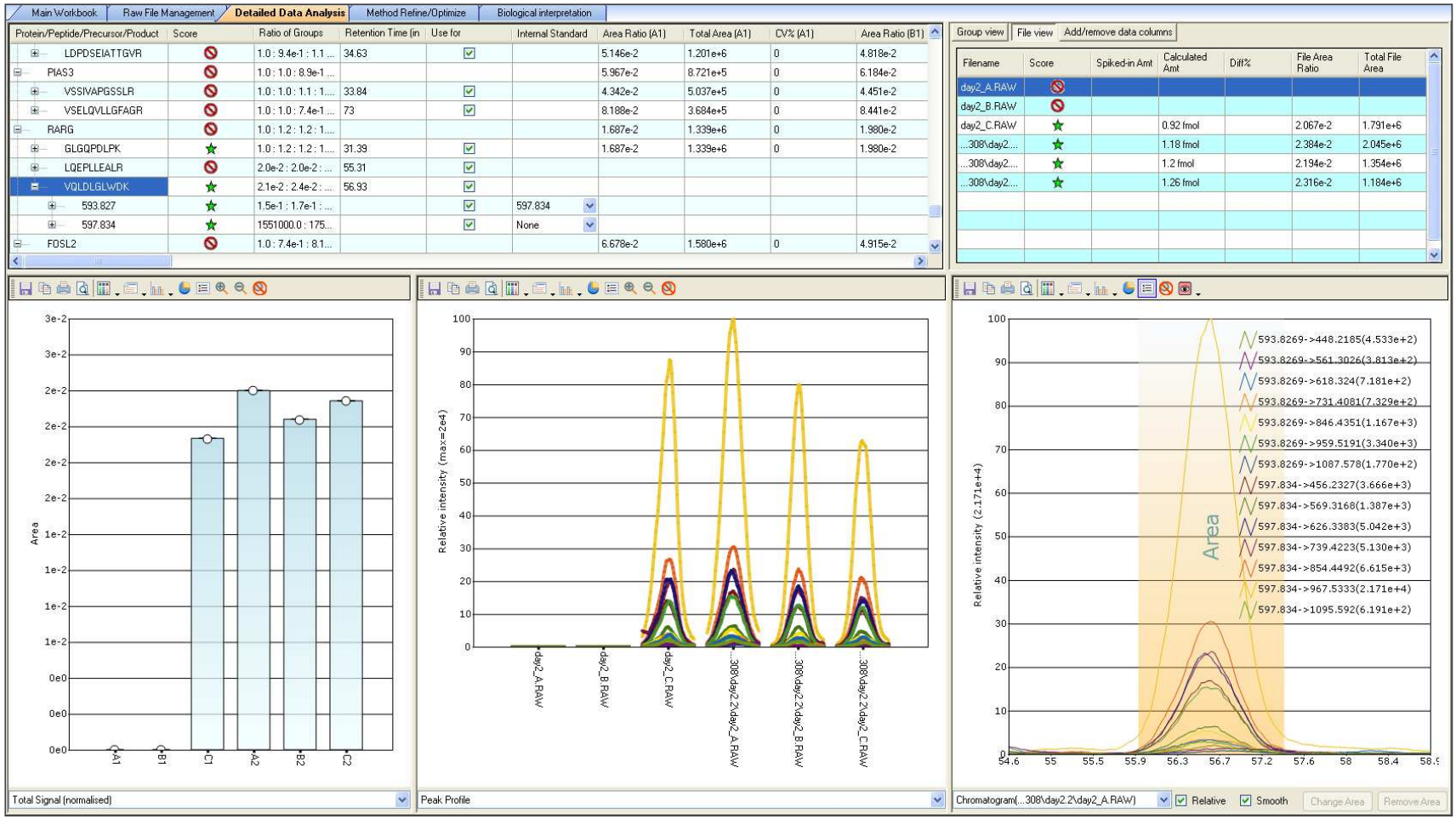

Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
737.879
LNHPESSQLFAK
Nfib
EDFVLTVTGK
GIPLESTDGER
587.291
592.295
Nr2C1
SPLAATPTFVTDSET...
AYVEFQDYITR
Smad2
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
943.87 fmol
1057.5 fmol
897.13 fmol
962.42 fmol
1171.38 fmol
990.97 fmol
Area
Relative intensity (max=9e4)
Relative intensity (9.187e+4)
587.2909->304.161(5.030e+2)
587.2909->361.1825(3.043e+4)
587.2909->476.2094(7.383e+3)
587.2909->577.2571(1.195e+4)
587.2909->664.2891(4.882e+4)
587.2909->793.3317(9.187e+4)
587.2909->906.4158(3.164e+4)
587.2909->1003.469(4.168e+4)
592.295->314.1693(5.473e+2)
592.295->371.1907(1.499e+3)
592.295->486.2177(3.021e+2)
592.295->587.2654(8.564e+2)
592.295->674.2974(2.185e+3)
592.295->803.34(3.873e+3)
592.295->916.424(1.458e+3)
592.295->1013.477(2.095e+3)
Total Signal (normalised)
Peak Profile
Chromatogram(day2_A.RAW)
Relative
Smooth
Change Area
Remove Area

f

SMAD2

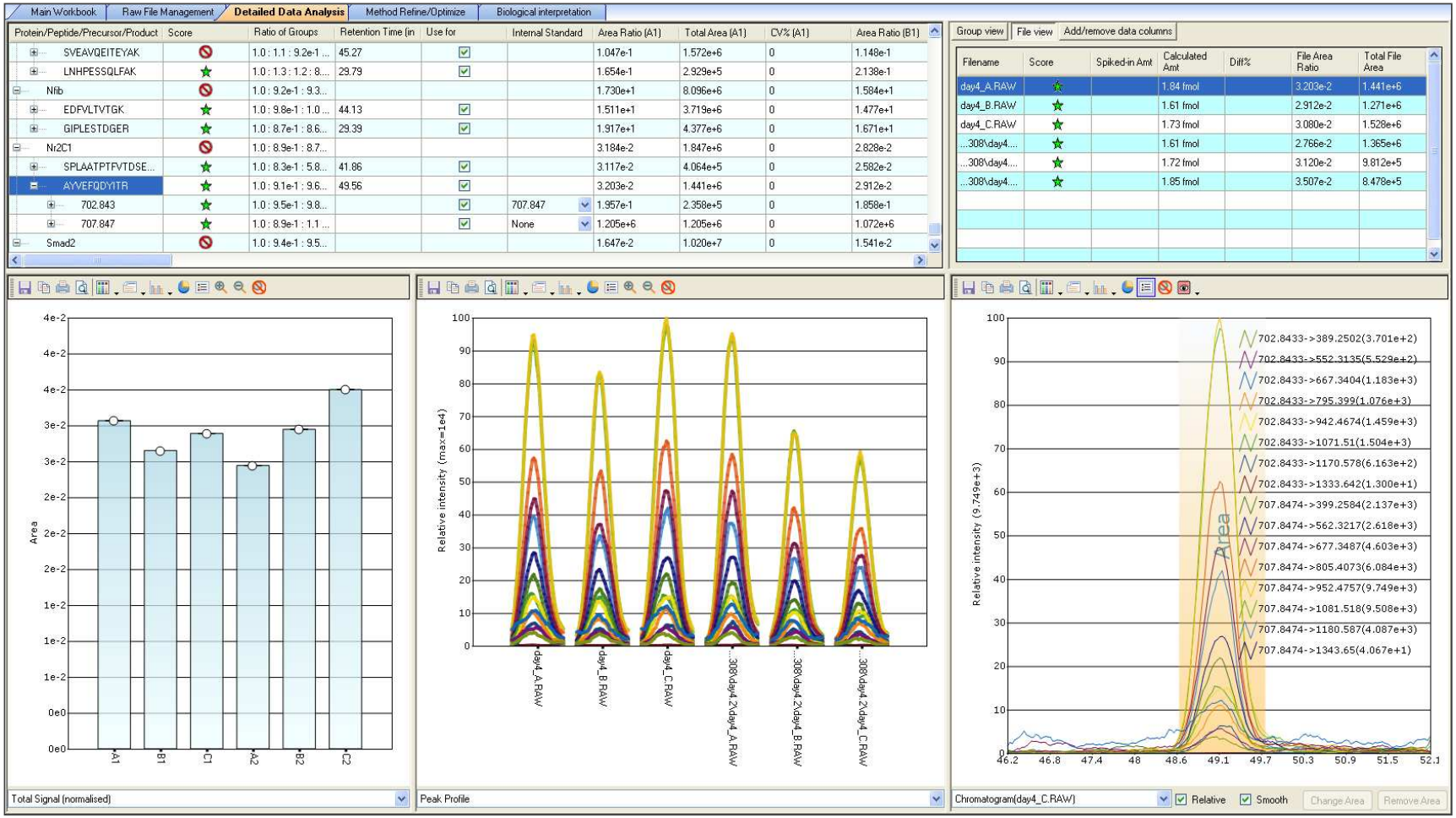

Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
GIPLESTDGER
587.291
592.295
Nr2C1
SPLAATPTFVTDSET...
AYVEFQDYITR
Smad2
VETPVLPPVLVPR
708.435
713.439
GWGAEYR
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
0.89 fmol
0.96 fmol
1.13 fmol
1.12 fmol
Area
Relative intensity (max=1e5)
Relative intensity (1.480e+5)
708.4346->371.2396(6.259e+2)
708.4346->583.392(3.247e+2)
708.4346->680.4448(1.619e+3)
708.4346->777.4976(2.229e+4)
708.4346->890.5817(3.711e+3)
708.4346->989.6501(1.154e+3)
708.4346->1086.703(3.788e+3)
713.4388->381.2479(5.174e+3)
713.4388->593.4003(2.382e+3)
713.4388->690.4531(8.249e+3)
713.4388->787.5059(1.480e+5)
713.4388->900.5899(2.358e+4)
713.4388->999.6583(6.794e+3)
713.4388->1096.711(2.610e+4)
Total Signal (normalised)
Peak Profile
Relative
Smooth
Change Area
Remove Area

g RARg

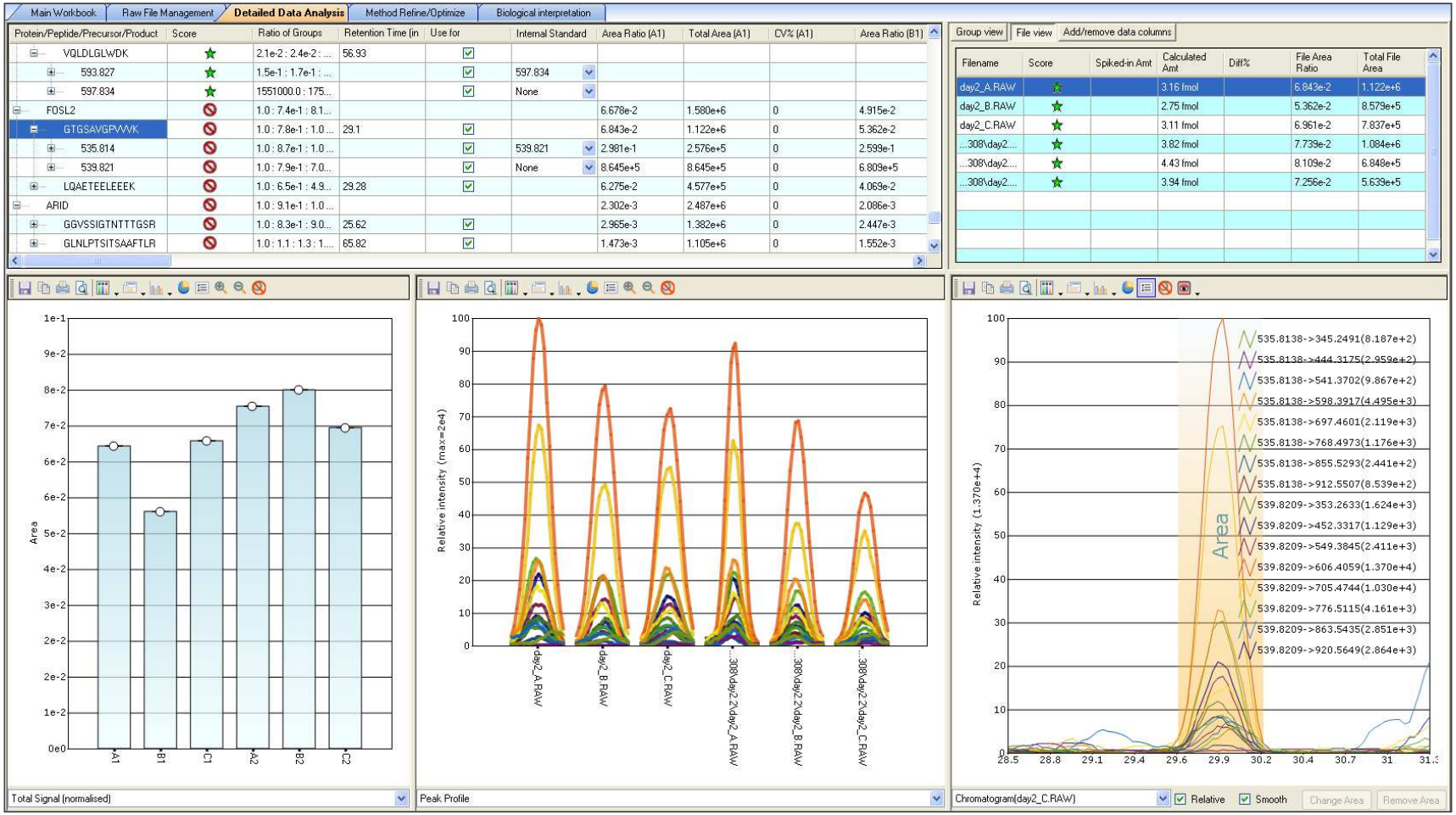

Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
LDPDSEIATTGVR
PIAS3
VSSIVAPGSSLR
VSELQVLLGFAGR
RARG
GLGQPDLPK
LQEPLLEALR
VQLDLGLWDK
593.827
597.834
FOSL2
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
0.92 fmol
1.18 fmol
1.2 fmol
1.26 fmol
Area
Relative intensity (max=2e4)
Relative intensity (2.171e+4)
593.8269->448.2185(4.533e+2)
593.8269->561.3026(3.813e+2)
593.8269->618.324(7.181e+2)
593.8269->731.4081(7.329e+2)
593.8269->846.4351(1.167e+3)
593.8269->959.5191(3.340e+3)
593.8269->1087.578(1.770e+2)
597.834->456.2327(3.666e+3)
597.834->569.3168(1.387e+3)
597.834->626.3383(5.042e+3)
597.834->739.4223(5.130e+3)
597.834->854.4492(6.615e+3)
597.834->967.5333(2.171e+4)
597.834->1095.592(6.191e+2)
Total Signal (normalised)
Peak Profile
Relative
Smooth
Change Area
Remove Area

h

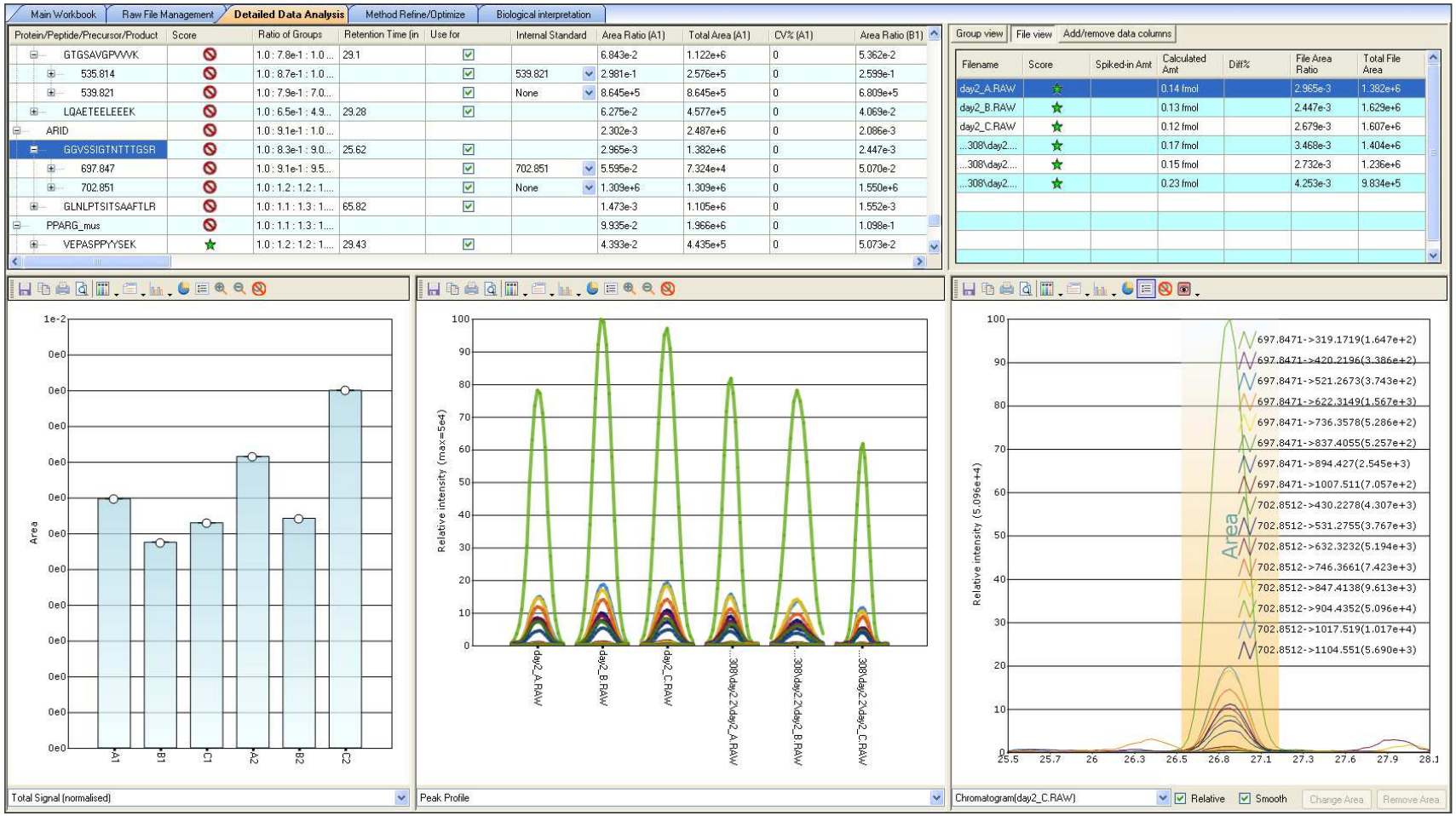
NR2C1


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
SVEAVQEITEYAK
LNHPESSQLFAK
Nfib
EDFVLTVTGK
GIPLESTDGER
Nr2C1
SPLAATPTFVTDSE...
AYVEFQDYITR
702.843
707.847
Smad2
Group view
File view
Add/remove data columns
Filename
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day4_A.RAW
day4_B.RAW
day4_C.RAW
Area
Relative intensity (max=1e4)
Relative intensity (9.749e+3)
Total Signal (normalised)
Peak Profile
Chromatogram(day4_C.RAW)
Relative
Smooth
Change Area
Remove Area

i FOSL2

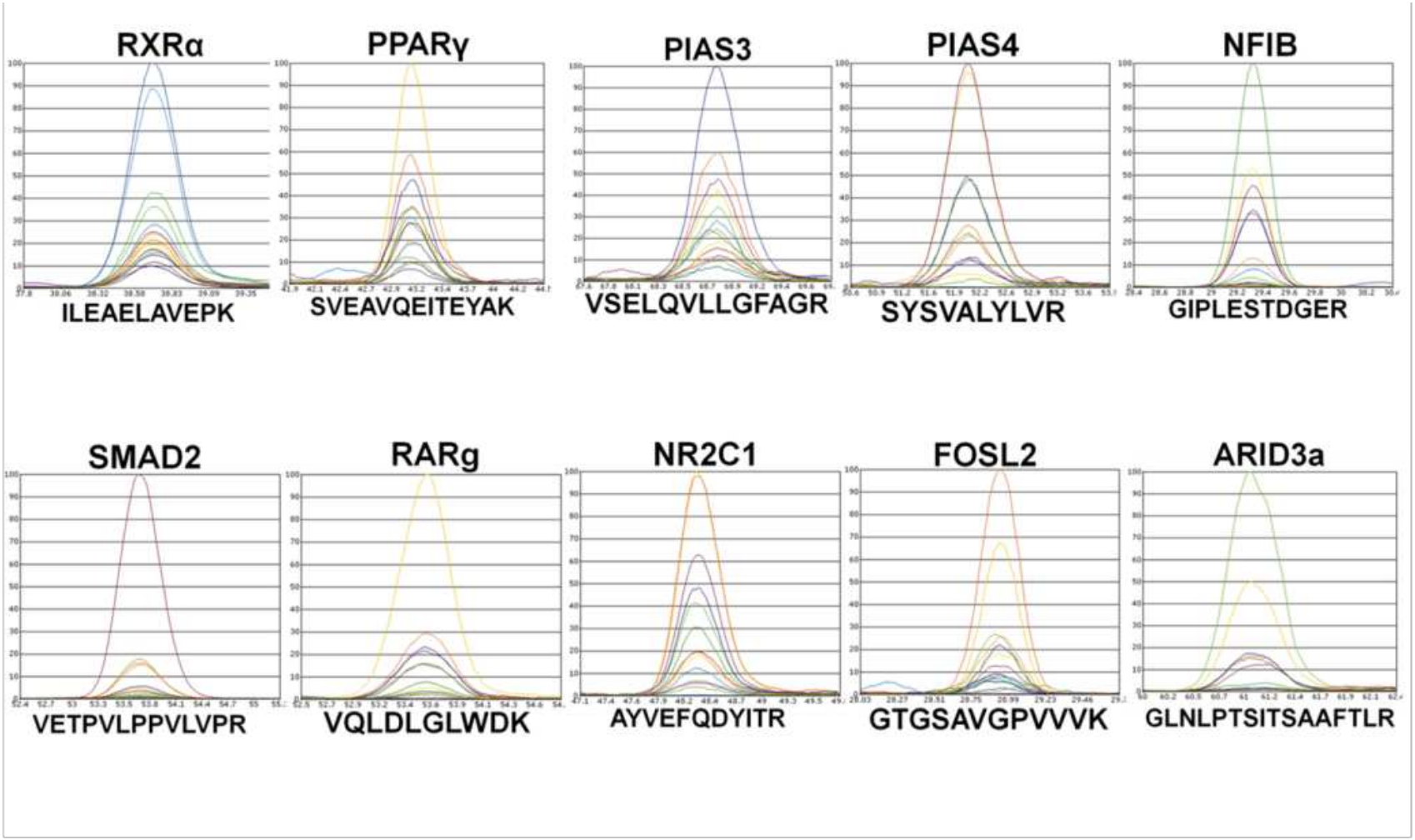

Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
VQLDLGLWDK
593.827
597.834
FOSL2
GTGSAVGPVVVK
535.814
539.821
LQAETEELEEEK
ARID
GGVSSIGTNTTTGSR
GLNLPTSITSAAFTLR
Group view
File view
Add/remove data columns
Filename
Score
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
Area
Relative intensity (max=2e4)
Relative intensity (1.370e+4)
Total Signal (normalised)
Peak Profile
Chromatogram(day2_C.RAW)
Relative
Smooth
Change Area
Remove Area

j ARID3a


Main Workbook
Raw File Management
Detailed Data Analysis
Method Refine/Optimize
Biological interpretation
Protein/Peptide/Precursor/Product
Score
Ratio of Groups
Retention Time (in
Use for
Internal Standard
Area Ratio (A1)
Total Area (A1)
CV% (A1)
Area Ratio (B1)
GTGSAVGPVVVK
535.814
539.821
LQAETEELEEEK
ARID
GGVSSIGTNTTTGSR
697.847
702.851
GLNLPTSITSAAFTLR
PPARG_mus
VEPASPPVYSEK
Group view
File view
Add/remove data columns
Filename
Score
Spiked-in Amt
Calculated Amt
Diff%
File Area Ratio
Total File Area
day2_A.RAW
day2_B.RAW
day2_C.RAW
Area
Relative intensity (max=5e4)
Relative intensity (5.096e+4)
Total Signal (normalised)
Peak Profile
Chromatogram(day2_C.RAW)
Relative
Smooth
Change Area
Remove Area

Supplementary Figure 17. (a-j) Screen prints of all peptides monitored by SRM for each of the 10 TFs at day two of differentiation.
The Pinpoint window provides a summary of the evaluation of the peptides of interest of the 10 TFs. Subpanels 1 and 2 show the calculated amount of peptide found for each technical replicate. Subpanel 3 shows the total normalized signal of each individual technical replicate. Subpanel 4 shows peak profiles of each technical replicate. Subpanel 5 shows the calculated peak area and chromatogram of peptide transitions. Panel legend and numbering applies to all figures.

Supplementary Figure 18



Supplementary Figure 18. Calculated peak areas of ten proteotypic peptides in their light and heavy versions.

Supplementary Tables

Supplementary Table 1

Supplementary Table 1: Comparison of available methods to our approach

Method	Pros (+) / Cons (-)	How our method compares	References
SILAC PrEST	+ Over 18,000 clones available for human. + Spiked early in the pipeline. + *In situ* quantification possible for one protein. + AB available for independent validation of the results. – External quantification needed in a separate step with intermediate storage step for each single protein selected for multiplexing. – Requires *in silico* selection of proteotypic peptides to select the right clone(s).	+ Full-length protein expression. + GST clones easily available. + *In situ* quantification possible. + No intermediate storage of protein needed (source of error). + External accurate quantification still possible (to extend the number of proteins quantified at a time). *Remark*: Total number of full-length protein or PrEST spikable is limited if absolute sensitivity is required, so no advantage in multiplexing too much.	Zeiler *et al.*, Mol Cell Proteomics. 2012 March; 11(3): O111.009613.
PSAQ & FLEXIQuant & Absolute SILAC (Approaches based on the use of full-length protein standard)	+ Mimics endogenous protein digestion behavior more accurately. + Spiked very early in the pipeline. – External quantification needed in a separate step with intermediate storage step for each single protein selected for multiplexing.	+ Multiplexing possible. + GST clones easily available. + Designed for multiplexed *in situ* quantification. + No intermediate storage of protein standard required (source of error).	Singh *et al.*, J Proteome Res. 2009 May; 8(5):2201-10. Brun *et al.*, J. Proteomics. 2009 Jul; 72(5):740-9. Hanke *et al.*, J. Proteome Res. 7, 1118–1130 (2008).
FLEXIQuant & EtEP (Approaches based on the use of a peptide-tag for quantification)	+ Simplification of the quantification procedure. + Only one peptide needs to be synthesized and accurately quantified. – *In situ* quantification limited to one protein (FlexiQuant) or loss of protein context (EtEP)	+ Multiplexing possible. + SH-Quant and the designed variants are proteotypic in mouse and human, stable over time in solution and sensitive. *Remark*: Peptidic tags are not equivalent. Therefore, any well-flying peptide is not necessarily a good candidate. Careful validation needed to ensure that peptide is quantotypic (digestability, stability in solution, efficient fragmentation). *Example*: Flexiquant is not stable over time in solution and less sensitive than SH-Quant.	Singh *et al.*, J Proteome Res. 2009 May; 8(5):2201-10. Holzmann *et al.*, Anal. Chem. 81, 10254–10261 (2009).
QconCAT	+ Highest multiplexing capacity among methods available. – Systematic optimization needed. – Loss of protein context. – Difficult to express. – Only useful if more than 20 proteins have to be quantified at a time.	+ Multiplexing. + Useful in methods centered on a limited number of (low abundant) proteins. + Relative amounts of protein spiked easily adaptable to the endogenous levels of protein targets.	Pratt *et al.*, Nat Protocol. 2006; 1(2):1029-43.

Supplementary Table 2

Supplementary Table 2: list of the 10 selected TFs

Gene symbol	Name	DNA-binding domain	IPI
RXRa	Retinoic acid receptor RXR-alpha	zf-C4	P28700
Nfib	Nuclear factor 1 B-type	MH1	P97863
Pias3	E3 SUMO-protein ligase PIAS3	SAP	O54714
Pias4	E3 SUMO-protein ligase PIAS4	SAP	Q9JM05
Fosl2	Fos-related antigen 2	bZIP_1	P47930
Rarg	Retinoic acid receptor gamma	zf-C4	P18911
PPARg	Peroxisome proliferator-activated receptor gamma	zf-C4	P37238
Arid3a	AT-rich interactive domain-containing protein 3A	ARID	Q62431
Nr2c1	Nuclear receptor subfamily 2 group C member 1	zf-C4	Q0VGP8
Smad2	Mothers against decapentaplegic homolog 2	MH1	Q62432

Supplementary Table 3

Supplementary Table 3: calculations of PPARγ and RXRα copy-number per cell (1 TF per assay)

		MS intensity	# cells	ug	ugNE/cell	copies/cell*
		1st biological replicate				
PPARγ	Day 0	0.039	21825000	2025	9.28E-05	2179
	2h	0.037	26475000	1876	7.09E-05	1579
	Day 1	0.028	36150000	2048	5.67E-05	955
	Day 2	0.069	39375000	2104	5.34E-05	2220
	Day 4	0.153	61125000	4227	6.92E-05	6372
	Day 6	0.072	56486250	3809	6.74E-05	2924
RXRα	Day 0	0.33	21825000	2025	9.28E-05	18438
	2h	0.252	26475000	1876	7.09E-05	10753
	Day 1	0.518	36150000	2048	5.67E-05	17672
	Day 2	1.399	39375000	2104	5.34E-05	45018
	Day 4	0.883	61125000	4227	6.92E-05	36772
	Day 6	0.357	56486250	3809	6.74E-05	14497
		2nd biological replicate				
PPARγ	Day 0	0.049	24975000	1380	5.53E-05	1630
	2h	0.04	24450000	1420	5.81E-05	1399
	Day 1	0.092	29287500	1532	5.23E-05	2898
	Day 2	0.281	42000000	2036	4.85E-05	8203
	Day 4	0.303	48000000	2819	5.87E-05	10716
	Day 6	0.299	43950000	3174	7.22E-05	13003
RXRα	Day 0	0.245	24975000	1380	5.53E-05	8152
	2h	0.314	24450000	1420	5.81E-05	10982
	Day 1	0.627	29287500	1532	5.23E-05	19751
	Day 2	0.945	42000000	2036	4.85E-05	27587
	Day 4	0.468	48000000	2819	5.87E-05	16552
	Day 6	0.281	43950000	3174	7.22E-05	12221
		3rd biological replicate				
PPARγ	Day 0	0.108	24900750	1270	5.10E-05	3317
	2h	0.074	29025000	1260	4.34E-05	1935
	Day 1	0.182	45234000	1500	3.32E-05	3634
	Day 2	0.342	53700000	2050	3.82E-05	7862
	Day 4	0.308	56212500	2500	4.45E-05	8249
	Day 6	0.235	57412500	2620	4.56E-05	6458
RXRα	Day 0	0.191	24900750	1270	5.10E-05	5866
	2h	0.321	29025000	1260	4.34E-05	8392
	Day 1	0.459	45234000	1500	3.32E-05	9166
	Day 2	1.099	53700000	2050	3.82E-05	25265
	Day 4	0.646	56212500	2500	4.45E-05	17301
	Day 6	0.445	57412500	2620	4.56E-05	12229

NE: nuclear extract

Calculations of 10 TFs copy-number per cell (multiplex)

		MS intensity	# cells	ug	ugNE/cell	copies/cell*
		4th biological replicate				
RXRα	Day 0	0.174	22650000	1264	5.58E-05	5847
	Day 2	0.551	20175000	1466	7.27E-05	24111
	Day 4	0.307	37275000	2249	6.03E-05	11155
PPARγ	Day 0	0.07	22650000	1264	5.58E-05	2352
	Day 2	0.194	20175000	1466	7.27E-05	8489
	Day 4	0.176	37275000	2249	6.03E-05	6395
Pias3	Day 0	0.023	22650000	1264	5.58E-05	773
	Day 2	0.011	20175000	1466	7.27E-05	481
	Day 4	0.007	37275000	2249	6.03E-05	254
Pias4	Day 0	0.172	22650000	1264	5.58E-05	5780
	Day 2	0.174	20175000	1466	7.27E-05	7614
	Day 4	0.136	37275000	2249	6.03E-05	4941
NFIB	Day 0	5.659	22650000	1264	5.58E-05	190178
	Day 2	6.984	20175000	1466	7.27E-05	305608
	Day 4	4.703	37275000	2249	6.03E-05	170879
SMAD2	Day 0	0.431	22650000	1264	5.58E-05	14484
	Day 2	0.341	20175000	1466	7.27E-05	14922
	Day 4	0.32	37275000	2249	6.03E-05	11627
RAR	Day 0	0.129	22650000	1264	5.58E-05	4335
	Day 2	0.106	20175000	1466	7.27E-05	4638
	Day 4	0.104	37275000	2249	6.03E-05	3779
NR2C1	Day 0	0.043	22650000	1264	5.58E-05	1445
	Day 2	0.041	20175000	1466	7.27E-05	1794
	Day 4	0.037	37275000	2249	6.03E-05	1344
FOSL2	Day 0	0.077	22650000	1264	5.58E-05	2588
	Day 2	0.093	20175000	1466	7.27E-05	4070
	Day 4	0.073	37275000	2249	6.03E-05	2652
ARID3a	Day 0	0.015	22650000	1264	5.58E-05	504
	Day 2	0.013	20175000	1466	7.27E-05	569
	Day 4	0.013	37275000	2249	6.03E-05	472

Supplementary Table 4

Supplementary Table 4: primers used to generate SH-quant variants

Tag name (amino acid substitution)		nucleotide sequence (5' to 3')
SH24-pF3A-GST (I4G)	FW	atcgatggaaaaagcggccgatggcacaagtttgtacaaaaaagc
	RV	gctttttgtacaaacttgtgccatcggccgctttttccatcgat
SH25-pF3A-GST (I4A)	FW	atcgatggaaaaagcggccgatgccacaagtttgtacaaaaaagc
	RV	gctttttgtacaaacttgtggcatcggccgctttttccatcgatt
SH26-pF3A-GST (I4V)	FW	cgatggaaaaagcggccgatgtcacaagtttgtacaaaa
	RV	ttttgtacaaacttgtgacatcggccgctttttccatcg
SH27-pF3A-GST (I4F)	FW	cgatggaaaaagcggccgatttcacaagtttgtacaaaa
	RV	ttttgtacaaacttgtgaaatcggccgctttttccatcg
SH28-pF3A-GST (D3E)	FW	cgatggaaaaagcggccgagatcacaagtttgtacaaaa
	RV	ttttgtacaaacttgtgatctcggccgctttttccatcg
SH29-pF3A-GST (D3E- I4G)	FW	gatcgatggaaaaagcggccgagggcacaagtttgtacaaaaaagctg
	RV	cagctttttgtacaaacttgtgccctcggccgctttttccatcgatc
SH30-pF3A-GST (D3E-I4A)	FW	gatcgatggaaaaagcggccgaggccacaagtttgtacaaaaaagctg
	RV	cagctttttgtacaaacttgtggcctcggccgctttttccatcgatc
SH31-pF3A-GST (D3E-I4V)	FW	cgatcgatggaaaaagcggccgaggtcacaagtttgtacaaaaaag
	RV	ctttttgtacaaacttgtgacctcggccgctttttccatcgatcg
SH32-pF3A-GST (D3E-I4F)	FW	cgatcgatggaaaaagcggccgagttcacaagtttgtacaaaaaag
	RV	ctttttgtacaaacttgtgaactcggccgctttttccatcgatcg

Supplementary Note

Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics

Simicevic et al.

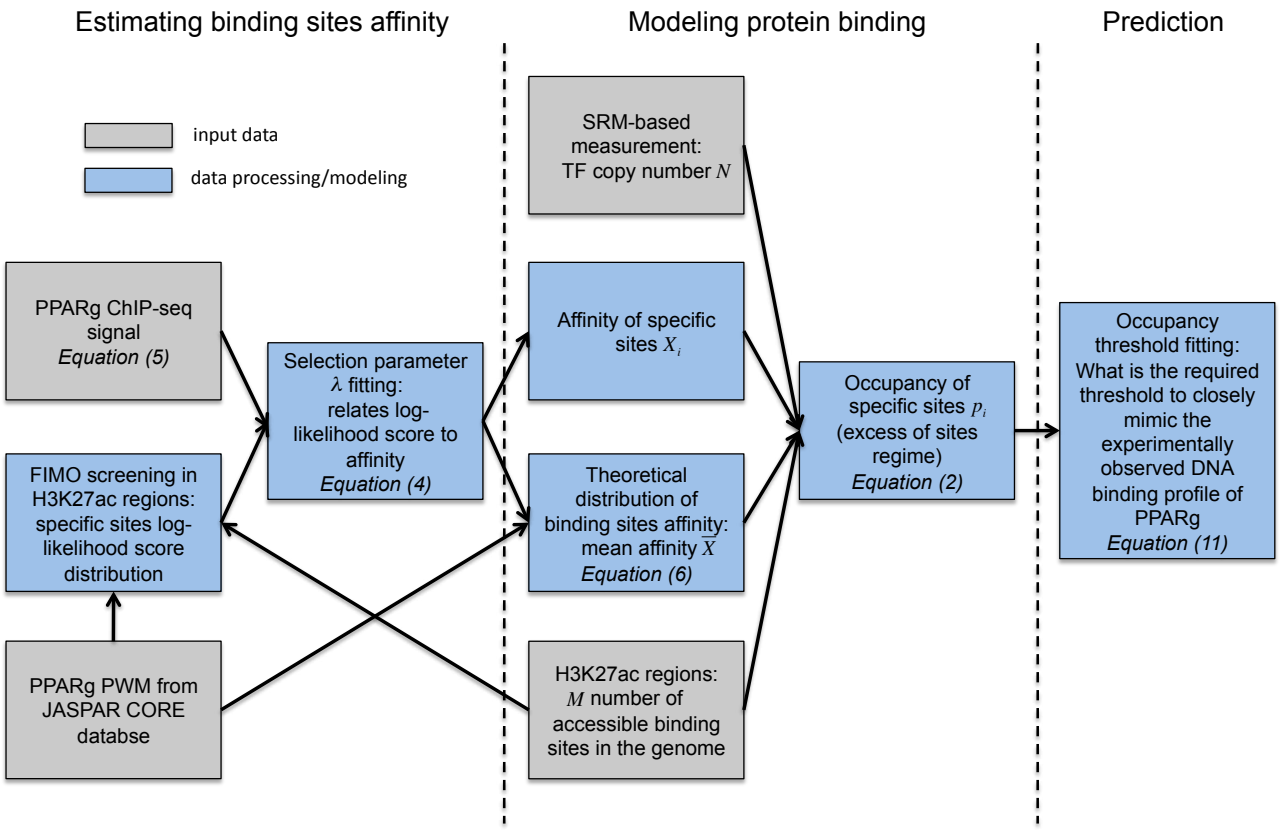
1 Outline

The goal of the following model is to determine if the temporal progression of detected PPAR γ binding sites during adipogenesis is consistent with the SRM-based measurement of PPAR γ copy numbers. The first part of this note addresses the modeling of protein binding (cf. Section 2 and Supplementary Fig. 19: Modeling protein binding) which aims at predicting the occupancy of the different specific sites p_i, assuming that we know the number of transcription factor molecules N, the size of the accessible genome M and the affinity of the different binding sites X_i. The second part of this note describes how to estimate the affinities X_i of the accessible specific sites from genome-wide ChIP-seq and accessibility data using the position weight matrix of the PPAR γ motif (cf. Section 3 and Supplementary Fig. 19: Estimating binding sites affinity). This allows us to reconstruct the affinity distribution of the different specific sites in the accessible regions of the genome. In the last part of this note, we estimate the occupancy threshold p_t above which PPAR γ binding sites are detected(cf. Section 4 and Supplementary Fig. 19: Prediction). We find that this threshold is in the range of 1% which is comparable to the % input of standard ChIP-seq experiments.

2 Statistical Modeling of protein binding

2.1 Statistical Mechanics Approach

The first question we want to address is how many protein molecules on average are bound to specific sites in a given cell. The approach used to compute the occupancy of the different specific sites is based on statistical mechanics. The main idea consists in enumerating all the different configurations by which one can place N proteins on the accessible genome weighted by the Boltzmann factor $\exp(-\beta E)$, which determines the likelihood of a configuration depending on its energy E. Typically, this is achieved by means of the partition function which describes the statistical properties of the system at thermodynamical equilibrium [1].



Supplementary Figure 19. Flowchart of the modelling process.

For the sake of simplicity, we made the following assumptions regarding our system: 1) we considered only one species of protein, 2) all the proteins are assumed to be on the DNA [2, 3] (either at specific sites or non-specific sites), 3) we do not consider hindrance between possibly (but very unlikely) overlapping sites, 4) we model k categories of specific sites of different affinity and all of these sites are assumed to be stronger than non-specific sites. Consequently, the system can be parametrized in the following way. N is the total number of proteins in the nucleus, M the size of the accessible genome, m_i the number of specific sites of category i, n_i the number of proteins bound to sites of category i and the energy E_i associated with each category of sites i. Therefore, the partition function Z of the system can be written as:

$$Z(N) = \sum_{n_1,n_2,\ldots,n_k=0}^{N} \binom{M-m}{N-n} \exp\left(-\beta(N-n)E_0\right) \left\{ \prod_{i=1}^{k} \binom{m_i}{n_i} \exp\left(-\beta n_i E_i\right) \right\}$$

with $m = \sum_{i=1}^{k} m_i$ the total number of specific sites, $n = \sum_{i=1}^{k} n_i$ the number of proteins bound to specific sites and E_0 the energy of an non-specific site. We can compute the average number of proteins $\bar{n}_i$ which are bound to specific sites of category i, with the derivative of the log of the

partition function with respect to the energy E_i:

$$\bar{n}_i = -\frac{1}{\beta}\frac{\partial}{\partial E_i}\log Z(N)$$

It follows that the mean occupancy p_i of a site i is given by:

$$p_i = \frac{\bar{n}_i}{m_i}$$

In the following, it is convenient to rewrite the partition function in term of the affinities $X_i = \exp\left(-\beta(E_i - E_0)\right)$ of the different sites where the reference energy is chosen as the energy of the non-specific sites E_0. Consequently, the affinity of a non-specific site is $X_0 = 1$ and the partition function is now given by:

$$Z(N) = \sum_{n_1,n_2,...,n_k=0}^{N} \binom{M-m}{N-n} \left\{ \prod_{i=1}^{k} \binom{m_i}{n_i} X_i^{n_i} \right\} \tag{1}$$

It is possible to compute an approximate expression for the average number of occupied specific sites $\bar{n}_i$ and their occupancy p_i in two different regimes: either the number of proteins is in excess compared to number of specific sites, or the number of specific sites is much larger than the number of proteins.

2.2 Regime 1: Excess of Proteins over Specific Sites

In this first regime, we assume that the number of proteins N is much larger than the total number of specific sites m but that N is still small compared to the size of the accessible genome M, namely:

$$M \gg N \gg m = \sum_{i=1}^{k} m_i \quad \text{with} \quad m_i \geq n_i \quad \forall i \in \{1, 2, ..., k\}$$

The average number of occupied sites and the mean occupancy of a site can be derived rigorously from the partition function by approximating the binomial coefficients. Here, we present a more intuitive approach leading to the same results. By setting $m_i = 1$ and $m_j = 0\,\forall j \neq i$ in the partition function (1), we can compute the probability $P(b|N, m = 1)$ that a specific site i is bound given that we have only one specific site in the genome and N proteins. We find that:

$$P(b|N, m = 1) = \frac{\frac{N}{M}X_i}{\frac{M-N}{M} + \frac{N}{M}X_i} \simeq \frac{\frac{N}{M}X_i}{1 + \frac{N}{M}X_i}$$

which is the famous Hill function where the occupation of the site depends on the concentration of proteins N/M and the affinity of the site X_i. Since there are many proteins compared to the number of specific sites, each site can be seen as independent from others, namely we neglect the depletion of the protein pool due to binding at specific sites. Therefore, the average number of specific sites which are occupied $\bar{n}_i$ is given by $P(b|N, m = 1)$ multiplied by the number of sites m_i:

$$\bar{n}_i = m_i \frac{\frac{N}{M}X_i}{1 + \frac{N}{M}X_i}$$

The mean occupancy $p_i = \bar{n}_i/m_i$ of a site of category i is then identical to $P(b|N, m = 1)$:

$$p_i = \frac{\frac{N}{M}X_i}{1 + \frac{N}{M}X_i}$$

2.3 Regime 2: Excess of Specific Sites over Proteins

In the second regime, we assume that the number of specific sites m_i and the number of non-specific sites $m_0 = M - m$ are much larger than the number of proteins N. This can be stated as:

$$M = \sum_{i=0}^{k} m_i \quad \text{with} \quad m_i \gg N \geq n_i \quad \forall i \in \{0, 1, ..., k\}$$

From the partition function (1), one can compute the probability $P(b|N = 1, \{m_i\})$ that one protein is bound to a specific site given the set of specific sites $\{m_i\}$ and that the total number of proteins is one, this probability is given by:

$$P(b|N = 1, \{m_i\}) = \frac{\frac{m_i}{M}X_i}{\frac{m_0}{M} + \sum_{i=1}^{k} \frac{m_i}{M}X_i}$$

The probability that one protein is bound to a specific site now depends on the concentrations of sites m_i/M and the affinities X_i of the different sites. Since the number of sites is much larger than the number of proteins, each protein can be considered as being independent from each other. The average number of proteins $\bar{n}_i$ which are bound to specific site i is then given by $P(b|N = 1, \{m_i\})$ times the number of proteins N:

$$\bar{n}_i = N\frac{\frac{m_i}{M}X_i}{\frac{m_0}{M} + \sum_{i=1}^{k} \frac{m_i}{M}X_i}$$

Thus, the expression for the mean occupancy $p_i = \bar{n}_i/m_i$ of a site of category i is similar that in the first regime except that the denominator now depends on the concentrations of sites m_i/M:

$$p_i = \frac{\frac{N}{M}X_i}{\frac{m_0}{M} + \sum_{i=1}^{k} \frac{m_i}{M}X_i}$$

In the above expression, since $X_0 = 1$, the denominator can be expressed in term of the average affinity of the sites $\bar{X}$ which depends on the distribution of affinities in the accessible genome $P(X_i) = m_i/M$, therefore we obtain:

$$p_i = \frac{\frac{N}{M}X_i}{\sum_{i=0}^{k} \frac{m_i}{M}X_i} = \frac{N}{M\bar{X}}X_i$$

2.4 Typical Parameters

The size of the accessible mice genome is typically less than 10% of the full genome, consequently $M \sim 10^8$ bp. Regarding the number of proteins, the measured number for $\mathrm{PPAR}\gamma$ and $\mathrm{RXR}\alpha$ were

approximately in the range of $N \sim 10^3$ proteins. The number of binding sites which were detected is approximately $m \sim 10^3$ sites [4], it is important to realize that those represent the stronger sites, weaker sites are most likely not detected. Finally, the dissociation constant K_d is typically in the order of $\mu\mathrm{M}$ for a non-specific site and nM for a strong site, consequently we can assume that a strong site has an affinity roughly 10^3 times larger than a non-specific one [3].

Given those numbers, the first regime (Section 2.2) where the proteins are assumed to be in excess is clearly not adequate, since the number of detected binding sites is in the range of the number of proteins 10^3. Despite the fact that in our case the strongest sites might be fewer than the number of proteins N, most of the sites will actually be in large excess compared to N, therefore the second regime (Section 2.3) is appropriate to describe the occupancies of the different sites in our ChIP-seq experiments. In fact, it also predicts very well the behavior of the strongest sites (cf. Supplementary Fig. 20), because the strongest sites are not saturated due to the competition with the very large number of less favorable sites. Consequently, we expect a linear relationship between the occupancy and the affinity of the sites.

$$p_i = \frac{N}{M\bar{X}} X_i \tag{2}$$

If the non-specific sites dominate, namely the affinities X_i decrease fast enough compared to the increase in number of sites m_i, then the mean affinity of the sites will be close to one $\bar{X} \simeq 1$. This might not be true in practice, we will see later how to estimate the coefficient $\bar{X}$ (Section 3.4).

3 Link between ChIP-seq signal and affinity

3.1 Simple ChIP-seq Model

In a ChIP-seq experiment, the signal, i.e. the average number of fragments S_i for any sites of type i, should reflect the mean occupancy of the sites p_i. Of course, the occupancy is not the only contribution to the signal, the efficiency and specificity of the antibody as well as the number of cells play an important role, but those effects can be implicitly included in a simple model:

$$S_i = Ap_i + B$$

where as a first approximation the signal S_i is proportional to the occupancy p_i plus some background B [5]. Since in our case the occupancy goes as $p_i \simeq \frac{N}{MX} X_i$, we expect a correlation between the affinity X_i and the mean signal S_i:

$$S_i = \frac{AN}{M\bar{X}} X_i + B \tag{3}$$

3.2 Link between Energy and PWM Log-likelihood

In practice, one does not necessarily know the binding energy of the sites of interest, but if we know the position weight matrix (PWM), it is then possible to make the link between each sequence and

its binding energy [5, 6]. In the PWM framework the binding energy of the site $\varepsilon = -\beta E$ is assumed to be the sum of independent contributions of each base $\varepsilon_k(b)$. If the binding sites are characterized by an average energy $\bar{\varepsilon}$, one can express the likelihood $q(s)$ that a randomly chosen binding sites of length L will have sequence s:

$$q(s) = \prod_{k=1}^{L} \frac{\exp\left(\lambda \varepsilon_k(s_k)\right)}{\sum_b \exp\left(\lambda \varepsilon_k(b)\right)} = \prod_{k=1}^{L} f_k(s_k)$$

where the selection parameter λ ensures that $\bar{\varepsilon} = \sum_s \varepsilon(s) q(s)$ and $f_k(b)$ is the frequency of observation of base b at position k. Therefore, the log-likelihood $z(s) = \log\left(q(s)\right)$ is related to energy by the factor λ plus a constant:

$$z = \log\left(q\right) = \lambda \varepsilon + \text{cst}$$

We can express the energy of the sites with respect to the consensus sequence which is assumed to be the best sequence with the largest log-likelihood $z_{max} = \log\left(q_{max}\right)$:

$$-\beta(E_i - E_c) = \frac{1}{\lambda}(z_i - z_{max})$$

where E_c is the energy of the consensus. Finally the affinities X_i in term of the log-likelihood are given by:

$$X_i = X_{max} \exp\left(\frac{1}{\lambda}(z_i - z_{max})\right) \tag{4}$$

where $X_{max} = \exp\left(-\beta(E_c - E_0)\right)$ is the affinity of the consensus sequence. In the following we will assume that $X_{max} = 10^3$ which is the typical magnitude for the strongest sites compared to non-specific ones [3].

3.3 Estimating the Parameter λ

Using our simple model for the ChIP-seq signal, we can now relate the log-likelihood score z_i of the different sites with the signal. Indeed, from equation (3) and (4), the signal is given by:

$$S_i = A' \exp\left(\frac{1}{\lambda} z_i\right) + B \tag{5}$$

where A' is a new proportionality constant. Therefore, it is possible to estimate the λ parameter of the $\mathrm{PPAR}\gamma$ motif from the ChIP-seq signal. This enables us to make the link between the energy scale and the log-likelihood score.

We focused on $\mathrm{PPAR}\gamma$ at day 0, day 2 and day 6, we used the motif from the JASPAR CORE database [7] and we screened the accessible genome with FIMO [8] from the MEME suite for all the sequences which had a p-value $< 10^{-3}$. Due to limited sequencing depth of available DHS data, we used the H3K27ac regions from Tarjei et al. [9] as a proxy for the accessible genome. We verified that these regions cover most of the reported $\mathrm{PPAR}\gamma$ sites. The size of these regions were similar between the three time points, $M_0 = 5.97 \cdot 10^7$ bp, $M_2 = 6.99 \cdot 10^7$ bp and $M_6 = 6.76 \cdot 10^7$ bp. For each of those three time points, we estimated the parameter λ using the mean ChIP-seq

signal for different category of sites, namely we separated the sequences in 6 bins of log-likelihood covering the range obtained with FIMO (cf. Supplementary Fig. 21). We estimated the background level for the three time points by taking a window of 200 bp shifted by 1000 bp from each peak, we obtained $B_0 = 1.273$, $B_2 = 1.146$ and $B_6 = 2.004$. Since there is not much ChIP signal at day 0 (very few $\mathrm{PPAR}\gamma$ proteins before induction), we estimated the λ parameter as the average of day 2 and 6, we obtained $\lambda \simeq 3.36$.

3.4 Estimating $\bar{X}$

In the second regime, where the specific sites are in excess compared to proteins, we showed in Section 2.3 that the occupancy can be expressed as follows:

$$p_i = \frac{\frac{N}{M} X_i}{\sum_{i=0}^{k} \frac{m_i}{M} X_i} = \frac{N}{M\bar{X}} X_i$$

where $\bar{X}$ is actually the average affinity of the sites $\bar{X} = \sum_{i=0}^{k} X_i P(X_i)$. We estimated $\bar{X}$ using the theoretical distribution of sites with respect to the log-likelihood $\rho(z)$ given by the PWM of $\mathrm{PPAR}\gamma$:

$$\bar{X} = \int_{z_{min}}^{z_{max}} X(z)\rho(z)\mathrm{d}z \tag{6}$$

where $X(z)$ is given by equation (4) above the non-specific threshold $z^\star = z_{max} - \lambda \log(X_{max})$ which corresponds to the limit where non-specific binding starts to dominate. Therefore, $X(z)$ can be expressed as:

$$X(z) = \begin{cases} X_{max} \exp\left(\frac{1}{\lambda}(z - z_{max})\right) & z > z^\star \\ X_0 = 1 & z \leq z^\star \end{cases} \tag{7}$$

We approximated the probability density distribution $\rho(z)$ as a normal distribution $\mathcal{N}(z; \bar{z}, \sigma_z)$. We sampled the distribution from the PWM in order to estimate the mean $\bar{z} = -32.69$ and the standard deviation $\sigma_z = 5.33$. Performing the integral (6), we obtained $\bar{X} \simeq 1.93$. A correction to the normal distribution based on the saddle-point approximation [10] did not change much the result ($\bar{X} \simeq 1.97$).

4 Predicting the number of binding sites

4.1 Occupancy Threshold

In order to predict the number of binding sites that should be detected in our ChIP-seq data, we assume that there is a certain occupancy threshold p_t above which we will start detecting the sites. We can now express this threshold in term of the log-likelihood. Indeed, from equation (2) and (4) the occupancy can be expressed as :

$$p_t = \frac{N X_{max}}{M\bar{X}} \exp\left(\frac{1}{\lambda}(z_t - z_{max})\right) \tag{8}$$

Inverting this relation gives the log-likelihood threshold z_t:

$$z_t = \lambda \log\left(\frac{M\bar{X}p_t}{NX_{max}}\right) + z_{max} \tag{9}$$

Given this threshold z_t and the density of sites with respect to the log-likelihood, we will be able to predict the number of binding sites $N_{sites}(z_t)$, by counting how many sites in the accessible genome (section 3.3) have a log-likelihood larger than z_t.

4.2 Number of Binding Sites

We used the same approach than the previous section 3.3 to screen the accessible genome in order to model the cumulative number of sites $N_{sites}(z)$ with respect to the log-likelihood z. In order to obtain a smooth representation, it is convenient to parametrize the tail of $N_{sites}(z)$ as a power-law:

$$N_{sites}(z) = C(z_0 - z)^k \tag{10}$$

where k is the power law exponent, C and z_0 are constants. We estimated the parameters C, k and z_0 for $\mathrm{PPAR}\gamma$ at day 0, day 2 and day 6 using non-linear least squares, this representation gives excellent results (cf. Supplementary Fig. 22). Knowing the log-likelihood threshold z_t above which we should start to detect sites, we can predict the number of sites by evaluating (10) in z_t:

$$N_{sites}(z_t) = C(z_0 - z_t)^k \quad \text{with} \quad z_t = \lambda \log\left(\frac{M\bar{X}p_t}{NX_{max}}\right) + z_{max} \tag{11}$$

We determined the log-likelihood threshold z_t for the three different time points using equation (9) and the measured proteins copy number $N_0 = 2376, N_2 = 6095, N_6 = 7462$ which were average over the three biological replicates. The occupancy threshold p_t is a free parameter which was assumed to be the same between each time points, it was chosen so that the squared error between the detected number of sites and the predicted one is minimized, we obtained $p_t \simeq 1.35\%$ (cf. Supplementary Fig. 23) which is comparable to % input values in a good ChIP experiment [11]. We remind the reader that % input would correspond to occupancy in the case of an ideal ChIP (100% efficient antibody).

4.3 Independent Validation

We tested our model with a different set of ChIP-seq data provided by Siersbaek et al. [12] although they only performed ChIP-seq on $\mathrm{PPAR}\gamma$ at day 2 and 6 after induction. Following a similar approach than in section 3.3 and using the same DNA accessibility marks, we fitted the selection parameter and obtained $\lambda \simeq 2.69$. Using the measured copy number as in section 4.2 and assuming no detected $\mathrm{PPAR}\gamma$ binding site at day 0, we then minimized the squared error between the detected and predicted number of binding sites for the three time points, and we obtained an occupancy threshold of $p_t \simeq 0.98\%$ (cf. Supplementary Fig. 23) which is in the same range as the value we obtained in the previous section (section 4.2). Despite the new selection parameter, the

prediction for both data set are in good agreement. If we average both the selection parameter and the number of detected binding sites from [4] and [12], the model is then in excellent agreement with the average detected binding sites (cf. Supplementary Fig. 23).

4.4 Discussion

Our model makes a few assumptions that merit discussions. First, since we did not know the affinity of the best sites *in vivo*, we assumed that the strongest sites, namely sites corresponding to the consensus sequence, were 1000 fold stronger than non-specific sites ($X_{max} = 10^3$), which is realistic [3]. In addition, we computed how the predicted occupancies p_i depend on this number trough the ratio $X_i/\bar{X}$ (cf. Supplementary Fig. 24). We observe two regimes: one when $X_{max} > 10^3$ where $X_i/\bar{X}$ becomes insensitive to X_{max}. Secondly when $X_{max} < 10^3$ (which is unlikely biologically), the predicted occupancies would decrease, but this could be compensated by lowering the occupancy threshold p_t. Thus, our conclusions on the behavior of the number of sites vs proteins number do not strongly depend on the presumed value of X_{max}.

As a second assumption, we did not model explicitly the dimer between $\mathrm{PPAR}\gamma$ and $\mathrm{RXR}\alpha$, which is justified since $\mathrm{PPAR}\gamma$ is in limiting amounts. Moreover, we neglected the hindrance due to other protein species which could distort the predicted occupancies. Indeed, many proteins are interacting with DNA and thereby obstruct the accessible sites [13]. In particular, this might reduce the size of the genomic regions that are effectively accessible for $\mathrm{PPAR}\gamma$ and consequently the occupancy of the sites might be higher. Nevertheless, the values derived for the occupancy (2) will not be strongly affected since the product $M\bar{X}$ would only be mildly changed if we assume that additional proteins would essentially deplete non-specific or weak $\mathrm{PPAR}\gamma$ binding sites. Indeed, if we assume that one third of the transcription factors of the mouse (~ 850 TFs) are expressed at an average level of 10^4 copies [14], we would obtain $N_{TF} \simeq 8.5 \cdot 10^6$ molecules in the nucleus which is still small compared to the measured accessible genome size $M \simeq 6.8 \cdot 10^7$.

References

[1] K. Huang. *Statistical Mechanics, 2nd Ed*. Wiley India Pvt. Limited, 2008.

[2] Y Kao-Huang, A Revzin, A P Butler, P O'Conner, D W Noble, and P H von Hippel. Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: measurement of DNA-bound Escherichia coli lac repressor in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 74(10):4228–4232, October 1977.

[3] Mark D Biggin. Animal Transcription Networks as Highly Connected, Quantitative Continua. *Developmental Cell*, 21(4):611–626, October 2011.

[4] R Nielsen, T A Pedersen, D Hagenbeek, P Moulos, R Siersbaek, E Megens, S Denissov, M Borgesen, K J Francoijs, S Mandrup, and H G Stunnenberg. Genome-wide profiling of PPARγ:RXR and RNA polymerase II occupancy reveals temporal activation of distinct

metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes & Development*, 22(21):2953–2967, November 2008.

[5] B C Foat, A V Morozov, and H J Bussemaker. Statistical mechanical modeling of genome-wide transcription factor occupancy data by MatrixREDUCE. *Bioinformatics*, 22(14):e141–e149, July 2006.

[6] O G Berg and P H von Hippel. Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters. *Journal of molecular biology*, 193(4):723–750, February 1987.

[7] J C Bryne, E Valen, M H E Tang, T Marstrand, O Winther, I da Piedade, A Krogh, B Lenhard, and A Sandelin. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Research*, 36(Database):D102–D106, December 2007.

[8] C E Grant, T L Bailey, and W S Noble. FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27(7):1017–1018, March 2011.

[9] Tarjei S Mikkelsen, Zhao Xu, Xiaolan Zhang, Li Wang, Jeffrey M Gimble, Eric S Lander, and Evan D Rosen. Comparative Epigenomic Analysis of Murine and Human Adipogenesis. *Cell*, 143(1):156–169, October 2010.

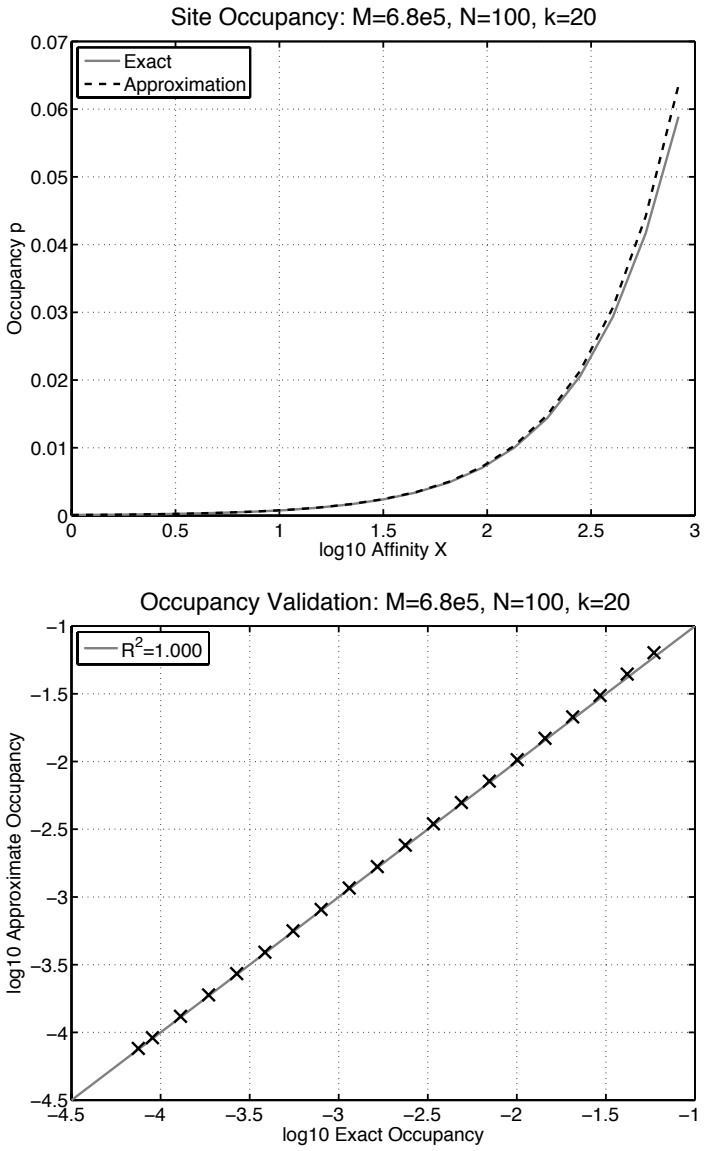
[10] M Djordjevic. A Biophysical Approach to Transcription Factor Binding Site Discovery. *Genome Research*, 13(11):2381–2390, November 2003.

[11] Guillaume Rey, François Cesbron, Jacques Rougemont, Hans Reinke, Michael Brunner, and Felix Naef. Genome-Wide and Phase-Specific DNA-Binding Rhythms of BMAL1 Control Circadian Output Functions in Mouse Liver. *PLoS Biology*, 9(2):e1000595, February 2011.

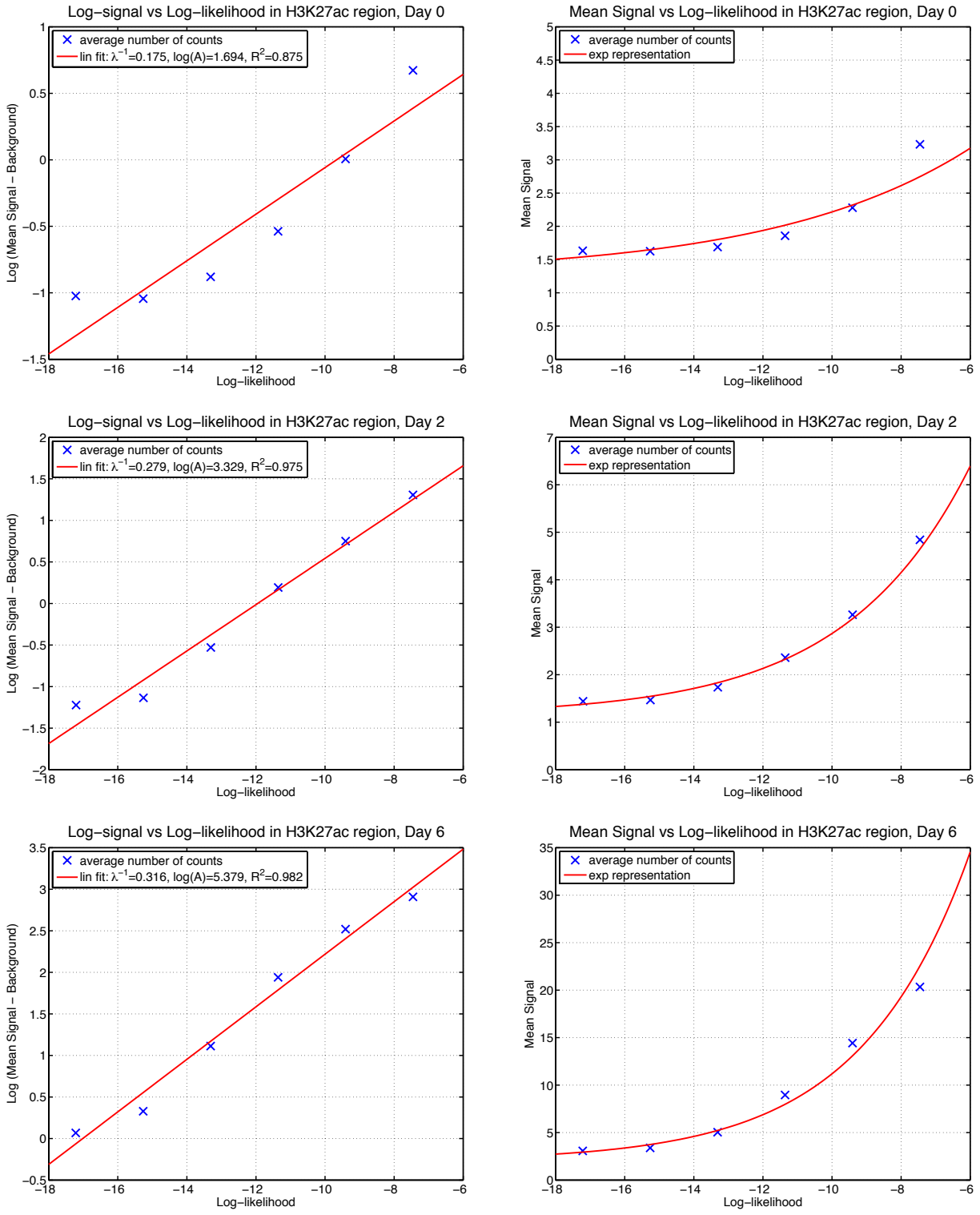
[12] Rasmus Siersbaek, Ronni Nielsen, Sam John, Myong-Hee Sung, Songjoon Baek, Anne Loft, Gordon L Hager, and Susanne Mandrup. Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. *The EMBO Journal*, 30(8):1459–1472, April 2011.

[13] Marco J Morelli, Rosalind J Allen, and Pieter Rein ten Wolde. Effects of Macromolecular Crowding on Genetic Networks. *Biophysj*, 101(12):2882–2891, December 2011.

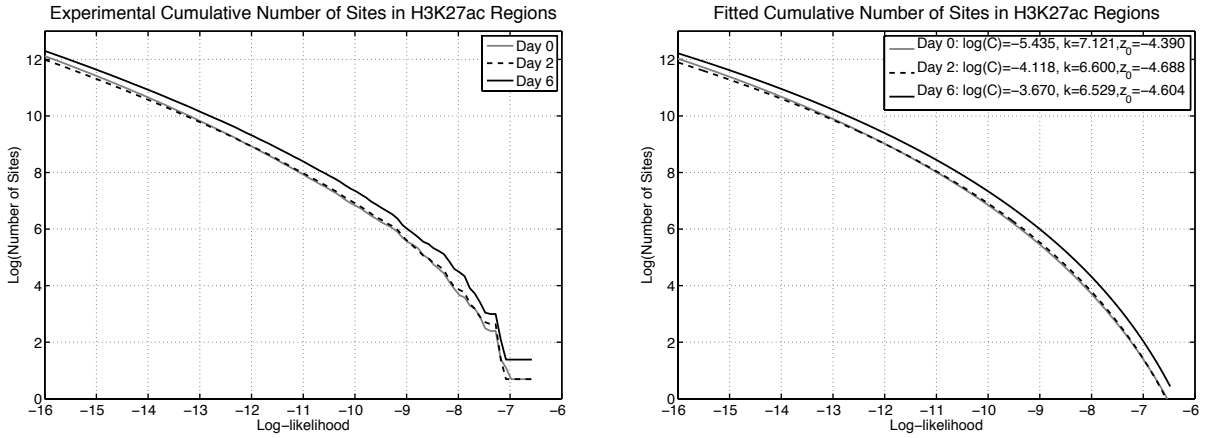
[14] S K Kummerfeld. DBD: a transcription factor prediction database. *Nucleic Acids Research*, 34(90001):D74–D81, January 2006.



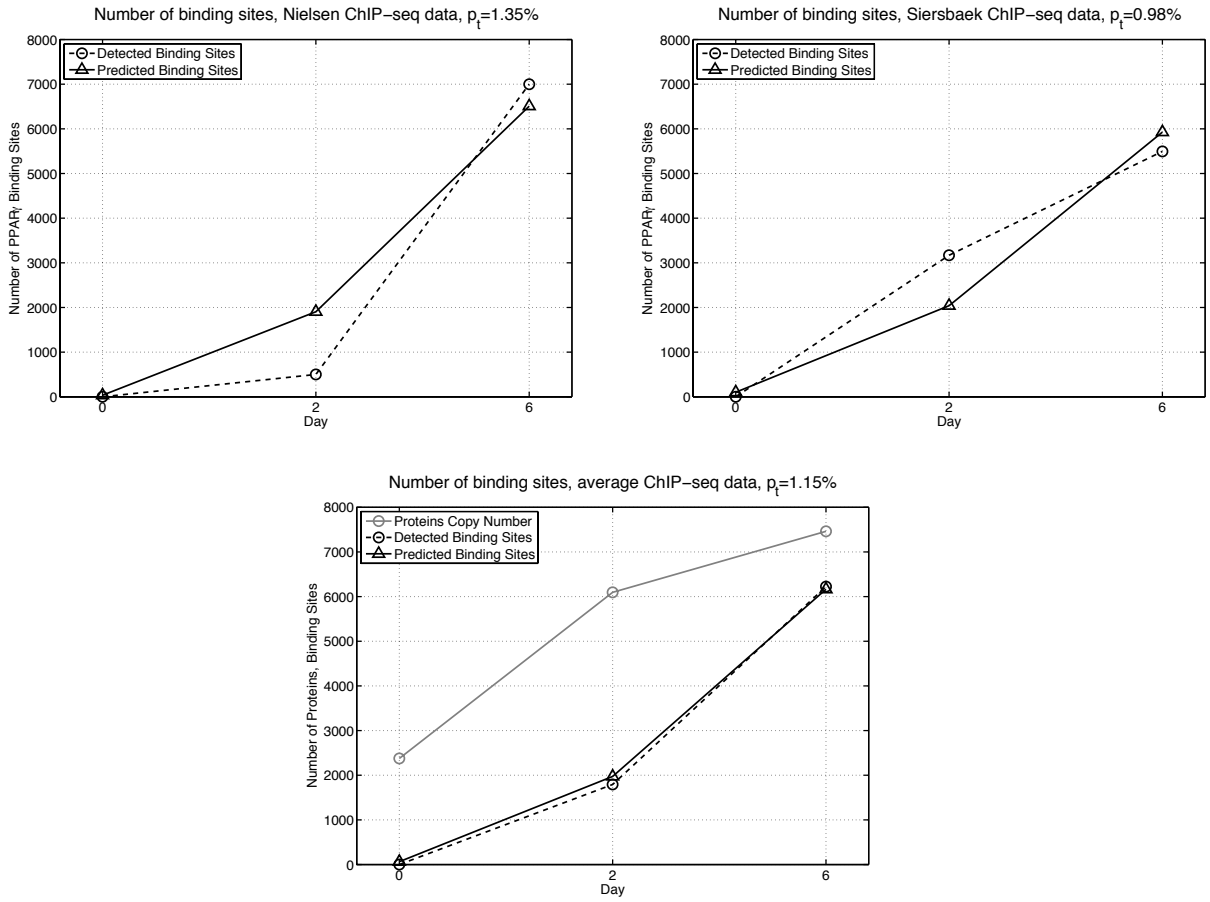
Supplementary Figure 20. Top: site occupancy p with respect to affinity X, computed with the approximate expression (2) and exactly from the partition function (1), for a reduced system with comparable concentration of proteins than $\mathrm{PPAR}\gamma$ at day 6 (cf. Section 3.3). The theoretical distribution of sites with respect to the affinity given by the PWM of $\mathrm{PPAR}\gamma$ was discretized in $k = 20$ categories, covering the range of affinities $[X_0 = 1, X_{max} = 1000]$. Bottom: comparison between the approximate and exact occupancy for the same system. Even if the stronger sites are in low amount compared to the number of proteins ($m_{19} = 1$, $m_{18} = 2$, ..., $m_{14} = 80 < N = 100$), the approximation (2) still predicts very well their occupancy.



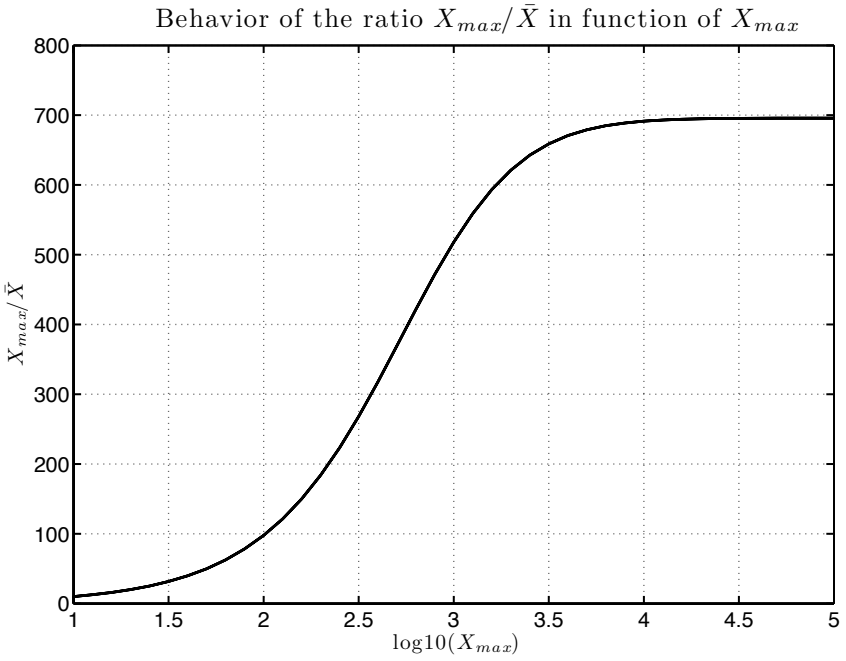
Supplementary Figure 21. Left column: fitting of the selection parameter λ from the mean ChIP-seq signal at day 0, day 2 and day 6 in log-space. Right column: representation of expression (5) in real space, using the estimated parameters on the left.



Supplementary Figure 22. Left: cumulative number of sites found in H3K27ac regions with FIMO [8] at day 0, day 2 and day 6. Right: cumulative number of sites approximated as a power law, this representation gives excellent results $R^2 > 0.99$.



Supplementary Figure 23. Left: predicted number of $\mathrm{PPAR}\gamma$ binding sites based on Nielsen et al. ChIP-seq data [4]. Right: predicted number of $\mathrm{PPAR}\gamma$ binding sites based on Siersbaek et al. ChIP-seq data. [12] Bottom: predicted number of $\mathrm{PPAR}\gamma$ binding sites compared to average number of detected sites in both [4] and [12].



Supplementary Figure 24. Effect of X_{max} on the ratio $X_{max}/\bar{X}$ which appears in the occupancy expression (8). If $X_{max} > 10^3$, the ratio becomes insensitive to X_{max}. On the other hand, if $X_{max} < 10^3$, the occupancy will decrease.

Supplementary Data 1: proteotypic peptide sele

PeptideAtlas query done on 14Apr2012

Biosequence Name
Peptide Accession
Pre AA
Sequence
Fol AA
Peptide Length
Combined Predictor Score
PSieve
ESPP
DPred
APEX
STEPP
N SP Mapping
N SP-varsplic Mapping
N SP-nsSNP Mapping
N ENSP Mapping
N ENSG Mapping
N IPI Mapping
N Human Mapping
N Mouse Mapping
N Yeast Mapping
Intensities
MQ.rank.MS1
TIC MS2 2+
TIC MS2 3+

Highlights PeptideAtlas Predictions:
Highlights PeptideAtlas Predictions:
Highlight Sequence

ction for the 10 TFs (multiplex)

Protein Name/Accession.
Peptide Atlas accession number, beginning with PAp followed by 9 digits.
Preceding (towards the N terminus) amino acid
Amino acid sequence of detected pepide, including any mass modifications.
Following (towards the C terminus) amino acid
Length of peptide
Score genereated based on STEPP, PSieve, ESPP, APEX and DPred scores
Predicted peptide score calculated by Peptide Sieve algorithm
Predicted peptide score calculated by ESPP algorithm
Predicted peptide score calculated by Detectability Predictor algorithm
Predicted peptide score calculated by APEX algorithm
Predicted peptide score calculated by STEPP algorithm
Number of SwissProt primary protein mapping
Number of SwissProt primary and alternatively-spliced protein mapping
Number of SwissProt primary and alternatively-spliced protein mapping, plus nsSNP mapping,wherein all Swiss-Prot-annotated nsSNPs have been expanded out to sequence with context so that any nsSNP-containing peptides are properly mapped.
Number of Ensembl protein mapping
Number of Ensembl gene mapping
Number of IPI protein mapping
Number of Human protein mapping, including SwissProt, IPI and Ensembl Proteins
Number of Mouse protein mapping, including SwissProt, IPI and Ensembl Proteins
Number of Yeast protein mapping, including SwissProt, SGD and Ensembl Proteins
MS1 intensities as determined using MaxQuant on OrbitrapXL data. Bar length proportional to intensity. The longer the higher.
MS1 intensity ranking. Bar length inversly proportional to ranking. The shorter the better.
Ranking of Total Ion Current in tandemMS spectrum as extracted from Scaffold for doubly charged ions. Bar length inversly proportional to ranking. The shorter the better.
Ranking of Total Ion Current in tandemMS spectrum as extracted from Scaffold for triply charged ions. Bar length inversly proportional to ranking. The shorter the better.

Number of protein mapped by peptide sequence between 1 and 3 were highlighted in green (Peptide sequence specificity)
Number of protein mapped by peptide sequence above 3 were not highlighted (Peptide sequence specificity)
Proteotypic peptides selected after filtering

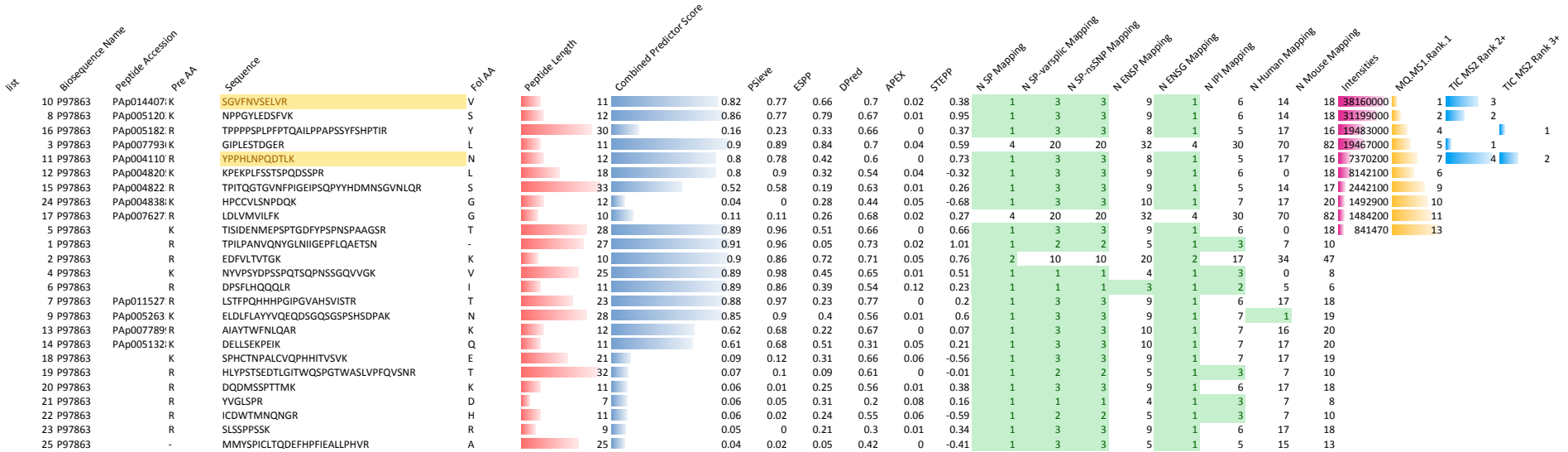
PeptideAtlas Legend

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities-y	MQ.MS1.Rank.2	TIC MS2 Rank 2+	TIC MS2 Rank 3+
1	P47930		R	GTGSAVGPVVVK	Q	12	0.92	0.83	0.66	0.78	0.1	0.61	1	1	1	1	1	2	0	4	3430300	1	1	
5	P47930	PAp014589	K	LQAETEELEEEK	S	12	0.78	0.75	0.37	0.6	0.02	0.57	1	1	1	1	1	2	6	4	2874200	2	2	
11	P47930		R	DEQLSPEEEEK	R	11	0.06	0.07	0.21	0.42	0	0.47	1	1	1	1	1	2	9	4	1007800	3		
9	P47930		R	SPPTSGLQSLR	G	11	0.18	0.18	0.63	0.69	0.02	0.33	1	1	1	1	1	2	0	4	918780	4		
13	P47930		K	EIAELQK	E	7	0.04	0.02	0.27	0.09	0	0.39	3	4	4	8	3	6	30	18	711070	5		
2	P47930	PAp017507	R	SSSSGDQSSDSLNSPTLLAL	-	20	0.92	0.91	0.48	0.82	0.04	0.81	1	1	1	1	1	2	7	4				
3	P47930	PAp001547	-	MYQDYPGNFDTSSR	G	14	0.85	0.93	0.33	0.61	0.01	0.33	1	1	1	1	1	2	3	4				
4	P47930	PAp000689	R	GSSGSPAHAESYSSGGGGQQK	F	21	0.85	0.85	0.36	0.78	0.01	0.34	1	1	1	1	1	2	3	4				
6	P47930		R	SHPYSPLPGLASVPGHMALPRPGVIK	T	26	0.73	0.8	0.22	0.68	0	-0.05	1	1	1	1	1	2	9	4				
7	P47930		K	QEPPEEDSPSSSAGMDK	T	17	0.72	0.67	0.25	0.62	0.01	0.76	1	1	1	1	1	2	0	4				
8	P47930	PAp005833	K	LEFMLVAHGPVCK	I	13	0.41	0.53	0.19	0.5	0.06	-0.19	1	1	1	1	1	1	9	3				
10	P47930		K	TIGTTVGR	R	8	0.07	0.02	0.38	0.67	0	0.28	1	1	1	1	1	2	9	4				
12	P47930		R	VDMPGSGSAFIPTINAITTSQDLQWMVQPTVITSMSNPYPR	S	41	0.04	0	0.1	0.44	0	-0.03	1	1	1	1	1	2	6	4				
14	P47930		K	ISPEER	R	6	0.03	0.01	0.11	0.08	0.01	0.05	2	4	4	8	3	6	9	18				

Proteotypic peptides

PeptideAtlas Predictions

Fosl2_P47930



list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities	MQ.MS1.Rank.1	TIC MS2 Rank 2+	TIC MS2 Rank 3+
10	P97863	PAp014407	K	SGVFNVSELVR	V	11	0.82	0.77	0.66	0.7	0.02	0.38	1	3	3	9	1	6	14	18	38160000	1	3	
8	P97863	PAp005120	K	NPPGYLEDSFVK	S	12	0.86	0.77	0.79	0.67	0.01	0.95	1	3	3	9	1	6	14	18	31199000	2	2	
16	P97863	PAp005182	R	TPPPPSPLPFPTQAILPPAPSSYFSHPTIR	Y	30	0.16	0.23	0.33	0.66	0	0.37	1	3	3	8	1	5	17	16	19483000	4		1
3	P97863	PAp007793	K	GIPLESTDGER	L	11	0.9	0.89	0.84	0.7	0.04	0.59	4	20	20	32	4	30	70	82	19467000	5	1	
11	P97863	PAp004110	R	YPPHLNPQDTLK	N	12	0.8	0.78	0.42	0.6	0	0.73	1	3	3	8	1	5	17	16	7370200	7	4	2
12	P97863	PAp004820	K	KPEKPLFSSTSPQDSSPR	L	18	0.8	0.9	0.32	0.54	0.04	-0.32	1	3	3	9	1	6	0	18	8142100	6		
15	P97863	PAp004822	R	TPITQGTGVNFPIGEIPSQPYYHDMNSGVNLQR	S	33	0.52	0.58	0.19	0.63	0.01	0.26	1	3	3	9	1	5	14	17	2442100	9		
24	P97863	PAp004838	K	HPCCVLSNPDQK	G	12	0.04	0	0.28	0.44	0.05	-0.68	1	3	3	10	1	7	17	20	1492900	10		
17	P97863	PAp007627	R	LDLVMVILFK	G	10	0.11	0.11	0.26	0.68	0.02	0.27	4	20	20	32	4	30	70	82	1484200	11		
5	P97863		K	TISIDENMEPSPTGDFYPSPNSPAAGSR	T	28	0.89	0.96	0.51	0.66	0	0.66	1	3	3	9	1	6	0	18	841470	13		
1	P97863		R	TPILPANVQNYGLNIIGEPFLQAETSN	-	27	0.91	0.96	0.05	0.73	0.02	1.01	1	2	2	5	1	3	7	10					
2	P97863		R	EDFVLTVTGK	K	10	0.9	0.86	0.72	0.71	0.05	0.76	2	10	10	20	2	17	34	47					
4	P97863		K	NYVPSYDPSSPQTSQPNSSGQVVGK	V	25	0.89	0.98	0.45	0.65	0.01	0.51	1	1	1	4	1	3	0	8					
6	P97863		R	DPSFLHQQQLR	I	11	0.89	0.86	0.39	0.54	0.12	0.23	1	1	1	3	1	2	5	6					
7	P97863	PAp011527	R	LSTFPQHHHPGIPGVAHSVISTR	T	23	0.88	0.97	0.23	0.77	0	0.2	1	3	3	9	1	6	17	18					
9	P97863	PAp005263	K	ELDLFLAYYVQEQDSGQSGSPSHSDPAK	N	28	0.85	0.9	0.4	0.56	0.01	0.6	1	3	3	9	1	7	1	19					
13	P97863	PAp007789	R	AIAYTWFNLQAR	K	12	0.62	0.68	0.22	0.67	0	0.07	1	3	3	10	1	7	16	20					
14	P97863	PAp005132	K	DELLSEKPEIK	Q	11	0.61	0.68	0.51	0.31	0.05	0.21	1	3	3	10	1	7	17	20					
18	P97863		K	SPHCTNPALCVQPHHITVSVK	E	21	0.09	0.12	0.31	0.66	0.06	-0.56	1	3	3	9	1	7	17	19					
19	P97863		R	HLYPSTSEDTLGITWQSPGTWASLVPFQVSNR	T	32	0.07	0.1	0.09	0.61	0	-0.01	1	2	2	5	1	3	7	10					
20	P97863		R	DQDMSSPTTMK	K	11	0.06	0.01	0.25	0.56	0.01	0.38	1	3	3	9	1	6	17	18					
21	P97863		R	YVGLSPR	D	7	0.06	0.05	0.31	0.2	0.08	0.16	1	1	1	4	1	3	7	8					
22	P97863		R	ICDWTMNQNGR	H	11	0.06	0.02	0.24	0.55	0.06	-0.59	1	2	2	5	1	3	7	10					
23	P97863		R	SLSSPPSSK	R	9	0.05	0	0.21	0.3	0.01	0.34	1	3	3	9	1	6	17	18					
25	P97863		-	MMYSPICLTQDEFHPFIEALLPHVR	A	25	0.04	0.02	0.05	0.42	0	-0.41	1	3	3	5	1	5	15	13					

Proteotypic peptides

PeptideAtlas Predictions

Nfib_P97863

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities.y	MQ.MS1.Rank.2	TIC MS2 Rank 2+	TIC MS2 Rank 3+
4	Q505F1		R	SPLAATPTFVTDSETAR	S	17	0.93	0.86	0.71	0.69	0.14	0.81	1	2	2	3	1	3	0	8	30690000	1	1	
6	Q505F1		K	AYVEFQDYITR	T	11	0.92	0.9	0.24	0.69	0.14	0.3	1	1	1	3	1	2	0	6	17958000	2	2	
26	Q505F1		R	AFDTLAK	A	7	0.06	0.02	0.29	0.36	0.04	0.47	2	2	2	5	2	4	20	11	8729600	3	3	
21	Q505F1		K	GLIGNVR	I	7	0.09	0.1	0.46	0.46	0.03	0.28	1	1	1	3	1	2	0	6	4016500	4	4	
3	Q505F1		K	EGPLLSESHVAFR	L	13	0.94	0.98	0.8	0.72	0.14	0.68	1	1	1	3	1	2	0	6	3445800	5		1
22	Q505F1		K	VFDLCVVCGDK	A	11	0.08	0	0.33	0.79	0.08	-0.59	1	2	2	3	1	3	9	8	2228400	6	5	
32	Q505F1		K	NLVYSCR	G	7	0.04	0.01	0.24	0.41	0.02	-0.64	1	2	2	3	1	3	11	8	1975700	7		
1	Q505F1		K	QFILANHEGSTPGK	V	14	0.95	0.98	0.78	0.69	0.17	0.95	1	2	2	3	1	3	0	8	1013700	8		
19	Q505F1		R	IDSVIPHILK	M	10	0.44	0.3	0.43	0.45	0.14	0.56	1	1	1	3	1	2	5	6	486940	9		
23	Q505F1		R	TYPDDTYR	L	8	0.07	0.03	0.25	0.59	0.03	0.15	1	1	1	3	1	2	5	6	323440	10		
24	Q505F1		K	LQEFCNSMVK	L	10	0.07	0.05	0.25	0.46	0.09	-0.46	1	1	1	3	1	2	4	6	111760	11		
2	Q505F1		K	IQIVTALDHSTQGK	Q	14	0.94	0.98	0.67	0.78	0.13	0.82	1	2	2	3	1	3	0	8				
5	Q505F1		R	SAGLLDSGMFVNIHPSGIK	T	19	0.92	0.97	0.23	0.74	0.08	0.61	1	1	1	3	1	2	0	6				
7	Q505F1		K	AIVLFSPDHPGLENMELIER	F	20	0.9	0.92	0.05	0.67	0.04	0.99	1	1	1	3	1	2	0	6				
8	Q505F1		K	DLSHCGGDMPVVQSLR	N	16	0.9	0.93	0.6	0.7	0.1	-0.32	1	1	1	3	1	2	0	6				
9	Q505F1		R	NGDTSFGAFHHDIQTNGDVSR	A	21	0.89	0.79	0.32	0.72	0.11	0.5	1	1	1	3	1	2	0	6				
10	Q505F1		R	LTMPSPMPEYLNVHYIGESASR	L	22	0.88	0.95	0.11	0.65	0.03	0.42	1	1	1	3	1	2	11	6				
11	Q505F1		K	VFLTTPDAAGVNQLFFTSPDLSAPHLQLLTEK	S	32	0.85	0.91	0.04	0.75	0	0.42	1	2	2	3	1	3	0	8				
12	Q505F1	PAp007731	R	LMNATITEELFFK	G	13	0.84	0.8	0.13	0.63	0.06	0.53	1	1	1	3	1	2	5	6				
13	Q505F1		K	ALTPGESSACQSPEEGMEGSPHLIAGEPSFVEK	E	33	0.72	0.76	0.62	0.57	0	0.11	1	1	1	3	1	1	0	5				
14	Q505F1		K	LCIDGHEYAYLK	A	12	0.64	0.55	0.16	0.6	0.14	-0.28	1	1	1	3	1	2	0	6				
15	Q505F1		K	TEPAMLMAPDK	A	11	0.63	0.6	0.65	0.46	0.02	0.69	1	1	1	3	1	2	0	6				
16	Q505F1		K	AESCQGDLSTLASVVTSLANLGK	A	23	0.6	0.67	0.2	0.71	0	-0.09	1	1	1	3	1	2	0	6				
17	Q505F1		K	MEPADYNSQIIGHSL	-	15	0.57	0.96	0.61	-1	0.05	1.64	1	1	1	3	1	2	0	6				
18	Q505F1		-	MATIEEIAHQIIDQQMGEIVTEQQTGQK	I	28	0.55	0.59	0.26	0.56	0.02	0.35	1	2	2	3	1	3	0	8				
20	Q505F1		K	SLMEHIFK	L	8	0.13	0.3	0.13	0.28	0.03	0.29	1	1	1	3	1	2	0	6				
25	Q505F1		R	HYGAITCEGCK	G	11	0.06	0	0.19	0.67	0.09	-0.77	1	2	2	3	1	3	0	8				
27	Q505F1		K	SSNCAASTEK	I	10	0.06	0.03	0.18	0.63	0.02	-0.24	1	2	2	3	1	3	9	8				
28	Q505F1		R	LLFLSMHWALSIPSFQALGQENSISLVK	A	28	0.05	0.01	0.02	0.65	0	-0.1	1	1	1	3	1	2	10	6				
29	Q505F1		K	SPDQGPNK	V	8	0.05	0	0.13	0.37	0.01	0.35	1	2	2	3	1	3	0	8				
30	Q505F1		K	QDSVQCER	K	8	0.05	0	0.21	0.64	0	-0.46	1	2	2	3	1	3	9	8				
31	Q505F1		R	CIAFGMK	Q	7	0.04	0.01	0.38	0.33	0.01	-0.23	1	2	2	3	1	3	9	8				
33	Q505F1		K	DCVINK	H	6	0.04	0	0.2	0.26	0.02	-0.37	1	2	2	3	1	3	0	8				
34	Q505F1		K	AYWNELFTLGLAQCWQVMNVATILATFVNCLHSSLQQDK	M	39	0.03	0	0.01	0.53	0	-0.98	1	1	1	3	1	2	0	6				
35	Q505F1		R	KPIEVSR	E	7	0.03	0.01	0.17	0.21	0.01	-0.46	1	2	2	3	1	3	9	8				

Proteotypic peptides

PeptideAtlas Predictions

Nr2c1_Q505F1

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities	MQ MS1 Rank 1	TIC MS2 Rank 2+	TIC MS2 Rank 3+
9	O54714	PAp004425	R	VSSIVAPGSSLR	E	12	0.88	0.87	0.83	0.77	0.01	0.34	1	3	3	4	1	4	0	11	3787000	1	2	
1	O54714	PAp015268	K	LTADPDSEVATTSLR	V	15	0.94	0.98	0.83	0.78	0.08	0.93	1	3	3	5	1	4	5	12	3032300	2	4	
7	O54714		R	FEEAHFTFALTPQQLQQILTSR	E	22	0.9	0.99	0.1	0.74	0.04	0.26	1	3	3	5	1	5	0	13	2822600	4		1
2	O54714		K	GALTSGHQPSSVLR	S	14	0.93	0.97	0.7	0.85	0.06	0.43	1	3	3	4	1	4	0	11	2537800	5	8	2
11	O54714	PAp015286	R	LSATVPNTIVVNWSSEFGR	N	19	0.85	0.83	0.08	0.77	0.05	0.33	1	3	3	5	1	5	9	13	2456100	6	3	4
6	O54714		R	NYSLSVYLVR	Q	10	0.9	0.8	0.37	0.67	0.17	0.09	1	3	3	5	1	5	5	13	1369700	7	5	
4	O54714	PAp015342	R	EGHGGPLPSGPSLTGCR	S	17	0.91	0.85	0.5	0.78	0.14	-0.04	1	3	3	4	1	4	5	11	805050	8		3
17	O54714		K	CDYTIQVQLR	F	10	0.14	0.06	0.36	0.65	0.15	-0.57	1	3	3	5	1	5	9	13	573440	9	7	
12	O54714		K	TLGPSDLSLLSLPPGTSPVGSPGPLAPIPPTLLTPGTLLGPK	R	42	0.84	0.92	0.35	0.58	0	0.37	1	2	2	3	1	3	0	8				5
3	O54714		R	VSELQVLLGFAGR	N	13	0.92	0.91	0.43	0.72	0.13	0.36	2	8	8	10	2	10	22	28			1	
24	O54714		K	ALHLLK	S	6	0.05	0.02	0.22	0.08	0.06	0.29	3	9	9	13	3	11	33	33			6	
30	O54714		R	LTVPCR	A	6	0.03	0	0.29	0.18	0.03	-0.48	2	4	4	6	2	5	5	15	2851300	3		
5	O54714		R	GTPSHFLGPLAPTLGSSHR	S	19	0.9	0.99	0.37	0.72	0.02	0.45	1	3	3	4	1	4	0	11				
8	O54714		K	HCPVTSAAIPALPGSK	G	16	0.88	0.82	0.78	0.61	0.11	0.03	1	3	3	4	1	4	0	11				
10	O54714		R	VEVIDLTIESSSDEEDLPPTK	K	21	0.88	0.95	0.24	0.49	0	1.19	1	3	3	4	1	4	5	11				
13	O54714		K	EASEVCPPPGYGLDGLQYSAVQEGIQPESK	K	30	0.82	0.91	0.45	0.53	0.01	0.18	1	3	3	4	1	4	0	11				
14	O54714		R	QLTAGTLLQK	L	10	0.78	0.6	0.68	0.66	0.09	0.42	1	3	3	5	1	5	5	13				
15	O54714	PAp015349	K	RPSRPINITPLAR	L	13	0.74	0.94	0.31	0.63	0	-0.97	1	3	3	5	1	5	9	13				
16	O54714		K	LCPLPGYLPPTK	N	12	0.22	0.23	0.45	0.39	0.13	0.01	1	3	3	5	1	5	9	13				
18	O54714		R	ALTCAHLQSFDAALYLQMNEK	K	21	0.08	0.24	0.08	0.54	0	-0.39	1	3	3	5	1	4	5	12				
19	O54714		R	SSTPAPPPGR	V	10	0.06	0.08	0.28	0.37	0	0.45	1	3	3	4	1	4	0	11				
20	O54714	PAp008570	-	MAELGELK	H	8	0.06	0.04	0.55	0.36	0	0.45	1	2	2	2	1	3	6	7				
21	O54714	PAp001381	K	KPTWTCPVCDK	K	11	0.06	0	0.29	0.66	0.09	-1.12	1	3	3	5	1	4	5	12				
22	O54714		R	EVLPGAK	C	7	0.05	0.01	0.21	0.27	0.03	0.57	2	5	5	6	2	7	13	18				
23	O54714		K	SSCAPSVQMK	I	10	0.05	0.01	0.33	0.39	0.04	-0.36	1	3	3	6	1	5	10	14				
25	O54714		R	VSLMCPLGK	M	9	0.05	0.01	0.49	0.24	0.04	-0.22	2	8	8	9	2	9	18	26				
26	O54714		R	FCLCETSCPQEDYFPPNLFVK	V	21	0.04	0	0.13	0.69	0.01	-0.82	1	3	3	5	1	5	9	13				
27	O54714		K	HELLAK	A	6	0.04	0.02	0.16	0.15	0.01	0.32	1	3	3	6	1	6	10	15				
28	O54714		K	HMVMSFR	V	7	0.04	0.01	0.19	0.17	0.04	-0.08	1	2	2	2	1	3	6	7				
29	O54714		K	NGAEPK	R	6	0.04	0	0.03	0.06	0.03	0.3	1	3	3	5	1	5	9	13				
31	O54714		R	NPDHSR	A	6	0.03	0	0.05	0.09	0.01	-0.02	3	9	9	12	3	11	23	32				
32	O54714		K	APYESLIIDGLFMEILNSCSDCDEIQFMEDGSWCPMKPK	K	39	0.03	0	0.05	0.37	0	-0.99	1	3	3	5	1	4	0	12				
33	O54714		R	SDVISLD	-	7	0.03	0.04	0.34	-1	0.06	1.11	1	3	3	4	1	4	0	11				
34	O54714		R	EVDMHPPLPQPVHPDVTMKPLPFYEVYGELIRPTTLASTSSQR	F	43	0.02	0.01	0.18	-1	0	-0.35	1	2	2	3	1	3							

Proteotypic peptides

PeptideAtlas Predictions

Pias3_O54714

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities	MQ MS1 Rank.1	TIC MS2 Rank 2+	TIC MS2 Rank 3+
20	Q9JM05	PAp020391	R	VSLICPLVK	M	9	0.06	0.05	0.37	0.31	0.07	-0.12	1	1	1	1	1	2	3	4	11113000	1	4	
11	Q9JM05	PAp020391	K	LQESPCIFALTPR	Q	13	0.5	0.49	0.65	0.52	0.09	-0.3	1	1	1	2	1	2	3	5	8903900	3	2	
10	Q9JM05	PAp020391	K	SYSVALYLVR	Q	10	0.52	0.36	0.38	0.68	0.13	0.09	1	1	1	3	1	3	3	7	8281900	4	1	
1	Q9JM05		R	TPLSGPTVDYPVLYGK	Y	16	0.94	0.97	0.58	0.8	0.12	0.75	1	1	1	2	1	2	0	5	7449100	5	6	
17	Q9JM05		K	AVQVVLR	I	7	0.07	0.03	0.34	0.47	0.04	0.18	1	1	1	2	1	2	3	5	7087600	7	5	
3	Q9JM05	PAp020391	R	ITVTWGNYGK	S	10	0.88	0.78	0.42	0.68	0.11	0.33	1	1	1	3	1	3	3	7	4887800	10	7	
27	Q9JM05		R	LSVPCR	A	6	0.04	0.01	0.24	0.13	0.06	-0.54	1	1	1	2	2	3	3	6	4345200	11	3	
16	Q9JM05		K	ILSECEGADEIEFLAEGSWRPIR	A	23	0.07	0.19	0.09	0.52	0.03	-0.68	1	1	1	1	1	1	0	3	3452300	13		1
15	Q9JM05	PAp010985	R	ICYSDTSCPQEDQYPPNIAVK	V	21	0.07	0.09	0.24	0.57	0.05	-0.45	1	1	1	2	1	2	3	5	607960	18		2
6	Q9JM05		R	QLTSSDLLQR	L	10	0.82	0.6	0.6	0.74	0.13	0.17	1	1	1	3	1	3	0	7	9916700	2		
7	Q9JM05	PAp005337	R	ALQLVQFDCSPELFK	K	15	0.78	0.66	0.2	0.54	0.16	-0.16	1	1	1	2	1	2	3	5	3462400	12		
19	Q9JM05		K	YLNGLGR	L	7	0.06	0.05	0.36	0.29	0.06	0.14	1	1	1	2	1	2	3	5	2070400	15		
5	Q9JM05	PAp007477	R	VSDLQMLLGFVGR	S	13	0.86	0.83	0.22	0.68	0.08	0.27	1	1	1	2	1	2	3	5	838390	16		
2	Q9JM05	PAp015426	R	LDPDSEIATTGVR	V	13	0.94	0.95	0.94	0.75	0.07	0.98	1	1	1	1	1	2	3	4				
4	Q9JM05		K	SAEPGPQAPRPLDPLALHSMPR	T	22	0.88	0.93	0.48	0.64	0.04	0.12	1	1	1	2	1	2	0	5				
8	Q9JM05		K	LPFFNMLDELLKPTELVPQSAEK	L	23	0.76	0.76	0.14	0.72	0.01	0.32	1	1	1	2	1	2	0	5				
9	Q9JM05	PAp005019	K	VNHSYCSVPGYYPSNKPGVEPK	R	22	0.66	0.7	0.34	0.46	0.09	-0.45	1	1	1	2	1	2	3	5				
12	Q9JM05		-	MAAELVEAK	N	9	0.19	0.21	0.58	0.4	0.06	0.56	1	1	1	2	1	2	3	5				
13	Q9JM05		K	TGADVVDLTLDSSSSSEDEDEDEDDDEDEDEGPRPK	R	36	0.13	0.24	0.34	0.43	0	0.62	1	1	1	1	1	1	0	3				
14	Q9JM05		K	GLVPAC	-	6	0.1	0.01	0.45	0.7	0.04	0.36	1	1	1	1	1	1	3	3				
18	Q9JM05		R	QVEMIR	N	6	0.06	0.02	0.29	0.68	0.01	-0.07	2	2	2	3	2	3	10	8				
21	Q9JM05		K	ELYETR	Y	6	0.06	0.01	0.09	0.7	0	-0.04	1	1	1	2	1	2	3	5				
22	Q9JM05		K	HELVTR	A	6	0.04	0.03	0.11	0.16	0.05	0.03	1	1	1	2	1	2	3	5				
23	Q9JM05		K	RPCRPINLTHLMYLSSATNR	I	20	0.04	0.08	0.18	0.73	0.01	-1.66	1	1	1	2	1	2	3	5				
24	Q9JM05		R	ELQPGVK	A	7	0.04	0.04	0.15	0.15	0	0.45	1	1	1	2	1	2	3	5				
25	Q9JM05	PAp020391	R	AETCAHLQCFDAVFYLQMNEK	K	21	0.04	0.03	0.05	0.47	0.04	-0.85	1	1	1	1	1	1	3	3				
26	Q9JM05		K	NMVMSFR	V	7	0.04	0.01	0.25	0.23	0.01	-0.05	1	1	1	2	1	2	3	5				
28	Q9JM05		K	KPTWMCPVCDKPAAYDQLIIDGLLSK	I	26	0.03	0	0.11	0.62	0.01	-1.38	1	1	1	1	1	1	0	3				
29	Q9JM05		K	TLKPEVR	L	7	0.03	0.02	0.22	0.2	0	-0.44	1	1	1	2	1	2	3	5				
30	Q9JM05		K	HPELCK	A	6	0.03	0.01	0.08	0.07	0.02	-0.4	1	1	1	3	1	3	3	7				

Proteotypic peptides

PeptideAtlas Predictions

Pias4_Q9JM05

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities.y	MQ.MS1.Rank.2	TIC.MS2.Rank.2+	TIC.MS2.Rank.3+
1	P37238	PAp00732	R	SVEAVQEITEYAK	N	13	0.95	0.97	0.72	0.81	0.13	0.91	1	2	2	2	1	3	11	7	24270000	1	1	
4	P37238	PAp00892	K	LLAEISSDIDQLNPESADLR	A	20	0.9	0.99	0.1	0.62	0	1.03	1	2	2	2	1	4	15	8	3060000	3	2	1
21	P37238		R	IFQGCQFR	S	8	0.05	0.02	0.26	0.48	0.03	-0.52	1	2	2	2	1	3	11	7	2658200	4	4	
7	P37238		K	VEPASPPYYSEK	T	12	0.84	0.67	0.51	0.61	0.1	0.71	1	2	2	2	1	4	20	8	2360500	5	3	
17	P37238		K	HLYDSYIK	S	8	0.08	0.1	0.13	0.26	0.09	0.26	1	2	2	2	1	4	16	8	2360400	6	5	
5	P37238	PAp01530	K	LNHPESSQLFAK	V	12	0.88	0.92	0.58	0.64	0.01	0.68	1	2	2	2	1	3	11	7	1463900	8	6	
9	P37238		K	NIPGFINLDLNDQVTLLK	Y	18	0.65	0.56	0.06	0.65	0.04	0.87	1	2	2	2	1	3	0	7	0	11	7	
15	P37238		K	HITPLQEQSK	E	10	0.15	0.11	0.41	0.56	0.07	0.49	1	2	2	2	1	3	11	7	3179100	2		
6	P37238		R	QIVTEHVQLLHVIK	K	14	0.85	0.71	0.27	0.63	0.13	0.43	1	2	2	2	1	3	0	7	2042400	7		
2	P37238	PAp01540	K	DGVLISEGQGFMTR	E	14	0.94	0.98	0.8	0.79	0.14	0.54	1	2	2	2	1	3	11	7	846270	9		
20	P37238		K	FEFAVK	F	6	0.05	0.04	0.35	0.3	0	0.37	2	3	3	3	2	4	24	10	829700	10		
3	P37238	PAp01537	K	TETDMSLHPLLQEIYK	D	16	0.91	0.91	0.25	0.68	0.07	0.95	1	2	2	2	1	3	11	7				
8	P37238		R	KPFGDFMEPK	F	10	0.68	0.62	0.46	0.56	0.08	-0.02	1	2	2	2	1	3	11	7				
10	P37238		R	ADPMVADYK	Y	9	0.46	0.41	0.59	0.65	0.03	0.59	1	2	2	2	1	4	0	8				
11	P37238		K	SPFVIYDMNSLMMGEDK	I	17	0.32	0.32	0.07	0.67	0.05	0.56	1	2	2	2	1	3	11	7				
12	P37238		K	TQLYNRPHEEPSNSLMAIECR	V	21	0.22	0.42	0.26	0.66	0.04	-0.92	1	2	2	2	1	4	0	8				
13	P37238		K	LQEYQSAIK	V	9	0.17	0.25	0.51	0.35	0.06	0.34	1	2	2	2	1	4	20	8				
14	P37238		K	ASGFHYGVHACEGCK	G	15	0.16	0.07	0.23	0.73	0.17	-0.74	4	5	5	7	4	7	38	19				
16	P37238		K	CLAVGMSHNAIR	F	12	0.09	0.14	0.42	0.69	0.02	-0.47	1	2	2	2	1	4	14	8				
18	P37238		K	YGVHEIIYTMLASLMNK	D	17	0.08	0.08	0.04	0.69	0	0.23	1	2	2	2	1	3	11	7				
19	P37238		R	AILTGK	T	6	0.05	0.02	0.33	0.26	0.02	0.42	3	4	4	4	3	9	25	17				
22	P37238		K	SFPLTK	A	6	0.04	0.01	0.18	0.2	0.03	0.31	1	2	2	2	1	5	17	9				
23	P37238		R	MPQAEK	E	6	0.03	0.01	0.07	0.06	0.01	0.16	3	6	6	5	3	9	15	20				
24	P37238		R	CDLNCR	I	6	0.03	0	0.12	0.16	0.03	-1.04	1	2	2	2	1	4	17	8				
25	P37238		K	FNALELDDSDLAIFIAVIILSGDRPGLLNVKPIEDIQDNLLQALELQLK	L	49	0.02	0	-1	-1	0	-0.02	1	2	2	2	1	3	11	7				

Proteotypic peptides

PeptideAtlas Predictions

Pparg_P37238

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities	MQ.MS1.Rank.1	TIC.MS2.Rank.2+	TIC.MS2.Rank.3+
16	P18911	PAp004090:	R	GLGQPDLPK	E	9	0.08	0.04	0.39	0.33	0.05	0.66	1	1	1	2	1	2	5	5	4498200	3	1	
9	P18911	PAp015285	K	LQEPLLEALR	L	10	0.81	0.72	0.6	0.51	0.08	0.42	1	3	3	2	1	3	10	8	7300400	1	2	
18	P18911		K	FSELATK	C	7	0.05	0.07	0.21	0.25	0.02	0.36	2	7	7	5	2	7	25	19	3495900	5	3	
12	P18911		R	VQLDLGLWDK	F	10	0.44	0.44	0.44	0.55	0.04	0.48	1	3	3	2	1	3	12	8	1161900	7	4	
1	P18911	PAp015404	R	LPGFTGLSIADQITLLK	A	17	0.93	0.98	0.18	0.76	0.1	0.6	1	3	3	2	1	3	13	8	378120	11	5	
15	P18911	PAp001645	K	SSGYHYGVSSCEGCK	G	15	0.13	0.05	0.24	0.71	0.16	-0.73	1	3	3	3	1	4	9	10	1233300	6	6	1
23	P18911		R	VYKPCFVCNDK	S	11	0.04	0	0.29	0.39	0.08	-1.18	1	3	3	3	1	4	10	10	3934600	4	7	2
6	P18911		R	LFAPGALGPGSGYPGAGFPFAFPGALR	G	27	0.86	0.86	0.27	0.87	0.02	0.19	1	1	1	2	1	2	0	5				3
19	P18911		K	IVEFAK	R	6	0.05	0.02	0.38	0.21	0	0.41	2	7	7	5	2	7	24	19	5391100	2		
17	P18911		K	AACLDILMLR	I	10	0.06	0	0.33	0.5	0.07	-0.35	1	3	3	2	1	3	13	8	452460	10		
2	P18911		R	YTPEQDTMTFSDGLTLNR	T	18	0.91	0.96	0.21	0.72	0.06	0.66	3	9	9	11	3	9	41	29				
3	P18911	PAp015400	K	EEGSPDSYELSPQLEELITK	V	20	0.9	0.96	0.37	0.63	0	1.01	1	3	3	3	1	4	12	10				
4	P18911	PAp008957	R	GSPPFEMLSPSFR	G	13	0.89	0.96	0.65	0.49	0.05	0.41	1	1	1	2	1	2	5	5				
5	P18911		K	AHQETFPSLCQLGK	Y	14	0.88	0.88	0.61	0.65	0.11	-0.26	2	7	7	6	2	8	25	21				
7	P18911	PAp004443	K	MEIPGPMPPLIR	E	12	0.84	0.74	0.49	0.54	0.1	0.44	1	3	3	2	1	3	10	8				
8	P18911		K	EMASLSVETQSTSSEEMVPSSPSPPPPPR	V	29	0.81	0.87	0.54	0.4	0.02	0.5	1	1	1	2	1	2	5	5				
10	P18911		R	EMLENPEMFEDDSSKPGPHPK	A	21	0.77	0.83	0.51	0.34	0.03	0.43	1	3	3	2	1	3	0	8				
11	P18911		K	ASSEDEAPGGQGK	R	13	0.54	0.52	0.16	0.58	0.04	0.61	1	3	3	2	1	3	0	8				
13	P18911		K	YTTNSSADHR	V	10	0.3	0.26	0.12	0.63	0.11	0.15	2	7	7	6	2	8	24	21				
14	P18911		R	RPSQPYMFPR	M	10	0.19	0.46	0.34	0.45	0.01	-0.58	1	3	3	2	1	3	10	8				
20	P18911		K	NMVYTCHR	D	8	0.04	0	0.15	0.55	0.04	-0.71	2	5	5	8	2	6	25	19				
21	P18911		K	CFEVGMSK	E	8	0.04	0.01	0.25	0.41	0.03	-0.41	2	7	7	6	2	8	38	21				
22	P18911		R	MDLEEPEK	V	8	0.04	0.02	0.34	0.13	0	0.37	1	3	3	2	1	3	10	8				
24	P18911		K	NCIINK	V	6	0.03	0	0.26	0.13	0.01	-0.33	2	5	5	9	2	6	25	20				
25	P18911		R	GQSPQPDQGP	-	10	0.02	0.01	0.28	-1	0.03	0.92	1	3	3	2	1	3	0	8				
26	P18911		R	TQMHNAGFGPLTDLVFAFAGQLLPLEMDDTETGLLSAICLICGDR	M	45	0.02	0	0.01	-1	0	-0.69	1	3	3	2	1	3	10	8				

Proteotypic peptides

PeptideAtlas Predictions

Rarg_P18911

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities	MQ.MS1.Rank.3	TIC.MS2.Rank.2+	TIC.MS2.Rank.3+	
7	P28700		K	HFLPLDFSTQVNSSSLNSPTGR	G	22	0.89	0.98	0.18	0.66	0.04	0.38	1	1	1	2	1	1	0	4	26781000	2	10	1	
6	P28700	PAp00747855	R	AGWNELLIASFSHR	S	14	0.9	0.94	0.26	0.52	0.11	0.2	3	4	4	12	3	8	39	24	20517000	3	4	4	
2	P28700	PAp00746849	K	ILEAELAVEPK	T	11	0.94	0.84	0.66	0.58	0.23	0.9	2	2	2	8	2	5	7	15	20343000	4	3		
3	P28700	PAp01449315	K	GLSNPAEVEALR	E	12	0.94	0.96	0.93	0.74	0.08	0.75	1	1	1	4	1	3	4	8	18740000	5	2		
10	P28700	PAp01926880	R	AIVLFNPDSK	G	10	0.79	0.69	0.85	0.34	0.07	0.7	1	1	1	4	1	3	4	8	16150000	7	1		
1	P28700	PAp00752323	R	NSAHSAGVGAIFDR	V	14	0.95	0.85	0.69	0.85	0.24	0.52	2	3	3	8	2	6	42	17	8635300	8	6	2	
18	P28700		R	VLTELVSK	M	8	0.08	0.04	0.41	0.48	0.03	0.42	3	4	4	12	3	8	45	24	8107900	10	5		
15	P28700	PAp01926875	K	VYASLEAYCK	H	10	0.08	0.05	0.38	0.52	0.09	-0.33	1	1	1	4	1	3	4	8	4544800	14	7		
16	P28700	PAp00029245	K	CLEHLFFFK	L	9	0.08	0	0.14	0.34	0.16	-0.26	3	4	4	12	3	8	39	24	3409500	16	9	5	
14	P28700	PAp00010394	K	HYGVYSCEGCK	G	11	0.08	0	0.17	0.66	0.14	-0.81	3	4	4	14	3	9	47	27	1128800	21	11		
8	P28700	PAp00606610	R	IPHFSELPLDDQVILLR	A	17	0.89	0.91	0.13	0.63	0.05	0.75	1	1	1	4	1	3	4	8	42668000	1	8	3	
13	P28700	PAp00745638	K	QLFTLVEWAK	R	10	0.36	0.36	0.18	0.71	0.05	0.28	3	4	4	12	3	8	46	24	17025000	6			
5	P28700	PAp01791977	K	VPAHPSGNMASFTK	H	14	0.92	0.97	0.6	0.69	0.04	0.69	1	1	1	5	1	4	4	10	8514500	9			
11	P28700		R	NENEVESTSSANEDMPVEK	I	19	0.7	0.7	0.37	0.47	0.01	0.78	1	1	1	4	1	3	0	8	6568800	11			
9	P28700	PAp01926882	K	DGILLATGLHVHR	N	13	0.87	0.89	0.43	0.58	0.05	0.45	3	4	4	12	3	8	39	24	5716300	13			
12	P28700	PAp01926887	K	TETYVEANMGLNPSSPNDPVTNICQAADK	Q	29	0.52	0.55	0.41	0.67	0.01	0.21	1	1	1	4	1	3	4	8	3176600	17			
19	P28700	PAp01926868	K	TELGCLR	A	7	0.04	0.02	0.3	0.5	0.01	-0.62	2	3	3	8	2	6	42	17	2345200	18			
23	P28700	PAp01638909	K	DLTYTCR	D	7	0.04	0.01	0.2	0.35	0.02	-0.67	1	1	1	5	1	4	4	10	1264100	20			
4	P28700		K	LIGDTPIDTFLMEMLEAPHQAT	-	22	0.93	0.99	0.05	0.75	0.04	1.43	1	1	1	4	1	3	0	8					
17	P28700	PAp01456379	K	HICAICGDR	S	9	0.08	0	0.26	0.57	0.15	-0.88	2	2	2	9	2	6	7	17					
20	P28700		R	EAVQEER	Q	7	0.04	0.01	0.18	0.29	0	0.11	3	4	4	12	3	8	47	24					
21	P28700	PAp01926863	K	YPEQPGR	F	7	0.04	0.02	0.23	0.26	0	0.1	2	2	2	8	2	5	7	15					
22	P28700		K	DCLIDK	R	6	0.04	0.01	0.24	0.16	0.04	-0.31	2	2	2	9	2	6	7	17					
24	P28700		R	DMQMDK	T	6	0.03	0	0.09	0.09	0.02	0.04	2	2	2	8	2	5	7	15					
25	P28700		K	CLAMGMK	R	7	0.03	0	0.36	0.13	0.01	-0.32	1	1	1	4	1	3	4	8					

Proteotypic peptides

PeptideAtlas Predictions

Rxra_P28700

list	Biosequence Name	Peptide Accession	Pre AA	Sequence	Fol AA	Peptide Length	Combined Predictor Score	PSieve	ESPP	DPred	APEX	STEPP	N SP Mapping	N SP-varsplic Mapping	N SP-nsSNP Mapping	N ENSP Mapping	N ENSG Mapping	N IPI Mapping	N Human Mapping	N Mouse Mapping	Intensities-y	MQ.MS1.Rank.2	TIC MS2 Rank 2+	TIC MS2 Rank 3+
1	Q62432	PAp004979	R	VGETFHASQPSLTVDGFTDPSNSER	F	25	0.94	0.94	0.51	0.7	0.19	0.7	1	2	2	4	1	2	8	8	490670	2		1
12	Q62432	PAp005012	R	YGWHPATVCK	I	10	0.08	0.11	0.21	0.6	0.05	-0.34	2	3	3	5	2	4	18	12	199240	4	3	
3	Q62432	PAp003856	R	VETPVLPPVLVPR	H	13	0.85	0.79	0.52	0.58	0.06	0.56	3	4	4	9	3	8	24	21	8815200	1	1	
8	Q62432	PAp000023	R	FCLGLLSNVNR	N	11	0.14	0.03	0.28	0.77	0.14	-0.39	5	6	6	13	5	13	37	32			2	
16	Q62432	PAp005281	K	*GWGAEYR*	R	7	0.05	0.05	0.21	0.36	0.04	0.03	2	3	3	5	2	4	18	12	367170	3		
2	Q62432	PAp004950	K	SAGGSGGAGGGEQNGQEEK	W	19	0.89	0.91	0.34	0.79	0.03	0.49	1	2	2	6	1	2	8	10				
4	Q62432	-		MSSILPFTPPVVK	R	13	0.58	0.62	0.55	0.43	0.02	0.47	1	2	2	6	1	2	8	10				
5	Q62432	K		AIENCEYAFNLK	K	12	0.57	0.61	0.41	0.51	0.07	-0.21	1	2	2	6	1	2	9	10				
6	Q62432	PAp015255	K	DEVCVNPYHYQR	V	12	0.36	0.41	0.19	0.46	0.13	-0.38	2	3	3	9	2	7	17	19				
7	Q62432	PAp004727	R	WPDLHSHHELK	A	11	0.22	0.37	0.19	0.31	0.06	0.3	1	2	2	6	1	2	9	10				
9	Q62432	PAp004321	K	GLPHVIYCR	L	9	0.1	0.07	0.32	0.65	0.08	-0.39	5	6	6	17	5	16	39	39				
10	Q62432	PAp001493	K	VLTQMGSPSVR	C	11	0.09	0.03	0.92	0.55	0.03	0.27	1	2	2	3	1	2	7	7				
11	Q62432	PAp003955	K	IFNNQEFAALLAQSVNQGFEAVYQLTR	M	27	0.09	0.08	0.04	0.75	0.01	0.12	2	3	3	5	2	4	18	12				
13	Q62432	R		NATVEMTR	R	8	0.07	0.06	0.53	0.6	0	0.13	1	2	2	4	1	2	8	8				
14	Q62432	R		QTVTSTPCWIELHLNGPLQWLDK	V	23	0.07	0.11	0.06	0.65	0.02	-0.29	2	3	3	4	2	4	17	11				
15	Q62432	PAp010515	K	AITTQNCNTK	C	10	0.06	0	0.28	0.6	0.04	-0.23	1	2	2	6	1	2	8	10				
17	Q62432	K		IPPGCNLK	I	8	0.05	0.01	0.5	0.32	0.04	-0.12	2	3	3	5	2	4	18	12				
18	Q62432	R		LDELEK	A	6	0.04	0.01	0.22	0.1	0.04	0.26	5	6	6	12	5	8	57	26				
19	Q62432	K		CVTIPSTCSEIWGLSTANTVDQWDTTGLYSFSEQTR	S	36	0.04	0	0.07	0.63	0	-0.66	1	1	1	3	1	1	0	5				
20	Q62432	R		LYYIGGEVFAECLSDSAIFVQSPNCNQR	Y	28	0.04	0	0.02	0.59	0	-0.86	2	3	3	5	2	4	18	12				
21	Q62432	R		LQVSHR	K	6	0.03	0.01	0.07	0.12	0.01	-0.11	5	6	6	17	5	16	42	39				

Proteotypic peptides

PeptideAtlas Predictions

Smad2_Q62432

Supplementary Data 2: precursor- to product-ion transitions selected for SRM (10 TFs)

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
ARID	GGVSSIGTNTTTGSR	697.8471	319.1719	y3
ARID	GGVSSIGTNTTTGSR	697.8471	420.2196	y4
ARID	GGVSSIGTNTTTGSR	697.8471	521.2673	y5
ARID	GGVSSIGTNTTTGSR	697.8471	622.3149	y6
ARID	GGVSSIGTNTTTGSR	697.8471	736.3578	y7
ARID	GGVSSIGTNTTTGSR	697.8471	837.4055	y8
ARID	GGVSSIGTNTTTGSR	697.8471	894.427	y9
ARID	GGVSSIGTNTTTGSR	697.8471	1007.511	y10
ARID	GGVSSIGTNTTTGSR	697.8471	1094.543	y11
ARID	GGVSSIGTNTTTGSR	697.8471	1181.575	y12
ARID	GGVSSIGTNTTTGSR	697.8471	1280.644	y13
ARID	GGVSSIGTNTTTGSR	697.8471	1337.665	y14
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	329.1802	y3
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	430.2278	y4
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	531.2755	y5
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	632.3232	y6
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	746.3661	y7
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	847.4138	y8
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	904.4352	y9
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	1017.519	y10
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	1104.551	y11
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	1191.583	y12
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	1290.652	y13
ARID	GGVSSIGTNTTTGSR[HeavyR]	702.8512	1347.673	y14
ARID	GLNLPTSITSAAFTLR	831.4646	389.2502	y3
ARID	GLNLPTSITSAAFTLR	831.4646	536.3185	y4
ARID	GLNLPTSITSAAFTLR	831.4646	607.3557	y5
ARID	GLNLPTSITSAAFTLR	831.4646	678.3928	y6
ARID	GLNLPTSITSAAFTLR	831.4646	765.4248	y7
ARID	GLNLPTSITSAAFTLR	831.4646	866.4725	y8
ARID	GLNLPTSITSAAFTLR	831.4646	979.5566	y9
ARID	GLNLPTSITSAAFTLR	831.4646	1066.589	y10
ARID	GLNLPTSITSAAFTLR	831.4646	1167.636	y11
ARID	GLNLPTSITSAAFTLR	831.4646	1264.689	y12
ARID	GLNLPTSITSAAFTLR	831.4646	1377.773	y13
ARID	GLNLPTSITSAAFTLR	831.4646	1491.816	y14
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	399.2584	y3
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	546.3268	y4
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	617.364	y5
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	688.4011	y6
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	775.4331	y7
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	876.4808	y8
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	989.5648	y9
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	1076.597	y10
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	1177.645	y11
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	1274.697	y12
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	1387.781	y13
ARID	GLNLPTSITSAAFTLR[HeavyR]	836.4688	1501.824	y14

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
FOSL2	GTGSAVGPVVVK	535.8138	345.2491	y3
FOSL2	GTGSAVGPVVVK	535.8138	444.3175	y4
FOSL2	GTGSAVGPVVVK	535.8138	541.3702	y5

FOSL2	GTGSAVGPVVVK	535.8138	598.3917	y6
FOSL2	GTGSAVGPVVVK	535.8138	697.4601	y7
FOSL2	GTGSAVGPVVVK	535.8138	768.4973	y8
FOSL2	GTGSAVGPVVVK	535.8138	855.5293	y9
FOSL2	GTGSAVGPVVVK	535.8138	912.5507	y10
FOSL2	GTGSAVGPVVVK	535.8138	1013.598	y11
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	353.2633	y3
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	452.3317	y4
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	549.3845	y5
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	606.4059	y6
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	705.4744	y7
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	776.5115	y8
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	863.5435	y9
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	920.5649	y10
FOSL2	GTGSAVGPVVVK[HeavyK]	539.8209	1021.613	y11
FOSL2	LQAETEELEEEK	724.3435	405.1974	y3
FOSL2	LQAETEELEEEK	724.3435	534.2401	y4
FOSL2	LQAETEELEEEK	724.3435	647.3241	y5
FOSL2	LQAETEELEEEK	724.3435	776.3667	y6
FOSL2	LQAETEELEEEK	724.3435	905.4093	y7
FOSL2	LQAETEELEEEK	724.3435	1006.457	y8
FOSL2	LQAETEELEEEK	724.3435	1135.5	y9
FOSL2	LQAETEELEEEK	724.3435	1206.537	y10
FOSL2	LQAETEELEEEK	724.3435	1334.595	y11
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	413.2116	y3
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	542.2542	y4
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	655.3383	y5
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	784.3809	y6
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	913.4235	y7
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	1014.471	y8
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	1143.514	y9
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	1214.551	y10
FOSL2	LQAETEELEEEK[HeavyK]	728.3506	1342.609	y11

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
NFIB	EDFVLTVTGK	554.7978	305.1814	y3
NFIB	EDFVLTVTGK	554.7978	404.2498	y4
NFIB	EDFVLTVTGK	554.7978	505.2975	y5
NFIB	EDFVLTVTGK	554.7978	618.3815	y6
NFIB	EDFVLTVTGK	554.7978	717.45	y7
NFIB	EDFVLTVTGK	554.7978	864.5184	y8
NFIB	EDFVLTVTGK	554.7978	979.5453	y9
NFIB	EDFVLTVTGK[HeavyK]	558.8049	313.1956	y3
NFIB	EDFVLTVTGK[HeavyK]	558.8049	412.264	y4
NFIB	EDFVLTVTGK[HeavyK]	558.8049	513.3117	y5
NFIB	EDFVLTVTGK[HeavyK]	558.8049	626.3958	y6
NFIB	EDFVLTVTGK[HeavyK]	558.8049	725.4642	y7
NFIB	EDFVLTVTGK[HeavyK]	558.8049	872.5326	y8
NFIB	EDFVLTVTGK[HeavyK]	558.8049	987.5595	y9
NFIB	GIPLESTDGER	587.2909	304.161	y2
NFIB	GIPLESTDGER	587.2909	361.1825	y3
NFIB	GIPLESTDGER	587.2909	476.2094	y4
NFIB	GIPLESTDGER	587.2909	577.2571	y5
NFIB	GIPLESTDGER	587.2909	664.2891	y6
NFIB	GIPLESTDGER	587.2909	793.3317	y7
NFIB	GIPLESTDGER	587.2909	906.4158	y8
NFIB	GIPLESTDGER	587.2909	1003.469	y9
NFIB	GIPLESTDGER[HeavyR]	592.295	314.1693	y2

NFIB	GIPLESTDGER[HeavyR]	592.295	371.1907	y3
NFIB	GIPLESTDGER[HeavyR]	592.295	486.2177	y4
NFIB	GIPLESTDGER[HeavyR]	592.295	587.2654	y5
NFIB	GIPLESTDGER[HeavyR]	592.295	674.2974	y6
NFIB	GIPLESTDGER[HeavyR]	592.295	803.34	y7
NFIB	GIPLESTDGER[HeavyR]	592.295	916.424	y8
NFIB	GIPLESTDGER[HeavyR]	592.295	1013.477	y9

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
NR2C1	AYVEFQDYITR	702.8433	389.2502	y3
NR2C1	AYVEFQDYITR	702.8433	552.3135	y4
NR2C1	AYVEFQDYITR	702.8433	667.3404	y5
NR2C1	AYVEFQDYITR	702.8433	795.399	y6
NR2C1	AYVEFQDYITR	702.8433	942.4674	y7
NR2C1	AYVEFQDYITR	702.8433	1071.51	y8
NR2C1	AYVEFQDYITR	702.8433	1170.578	y9
NR2C1	AYVEFQDYITR	702.8433	1333.642	y10
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	399.2584	y3
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	562.3217	y4
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	677.3487	y5
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	805.4073	y6
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	952.4757	y7
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	1081.518	y8
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	1180.587	y9
NR2C1	AYVEFQDYITR[HeavyR]	707.8474	1343.65	y10
NR2C1	SPLAATPTFVTDSETAR	882.4441	347.2032	y3
NR2C1	SPLAATPTFVTDSETAR	882.4441	476.2458	y4
NR2C1	SPLAATPTFVTDSETAR	882.4441	563.2778	y5
NR2C1	SPLAATPTFVTDSETAR	882.4441	678.3047	y6
NR2C1	SPLAATPTFVTDSETAR	882.4441	779.3524	y7
NR2C1	SPLAATPTFVTDSETAR	882.4441	878.4208	y8
NR2C1	SPLAATPTFVTDSETAR	882.4441	1223.59	y11
NR2C1	SPLAATPTFVTDSETAR	882.4441	1395.675	y13
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	573.2861	y5
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	789.3607	y7
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	888.4291	y8
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	1136.545	y10
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	1233.598	y11
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	1334.646	y12
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	1405.683	y13
NR2C1	SPLAATPTFVTDSETAR[HeavyR]	887.4482	1476.72	y14

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
PIAS3	VSSIVAPGSSLR	586.8353	375.2345	y3
PIAS3	VSSIVAPGSSLR	586.8353	462.2665	y4
PIAS3	VSSIVAPGSSLR	586.8353	519.288	y5
PIAS3	VSSIVAPGSSLR	586.8353	616.3408	y6
PIAS3	VSSIVAPGSSLR	586.8353	687.3779	y7
PIAS3	VSSIVAPGSSLR	586.8353	786.4463	y8
PIAS3	VSSIVAPGSSLR	586.8353	899.5303	y9
PIAS3	VSSIVAPGSSLR	586.8353	986.5624	y10
PIAS3	VSSIVAPGSSLR	586.8353	1073.594	y11
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	385.2428	y3
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	472.2748	y4
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	529.2963	y5
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	626.349	y6

PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	697.3861	y7
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	796.4545	y8
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	909.5386	y9
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	996.5706	y10
PIAS3	VSSIVAPGSSLR[HeavyR]	591.8394	1083.603	y11
PIAS3	VSELQVLLGFAGR	694.8984	303.177	y3
PIAS3	VSELQVLLGFAGR	694.8984	450.2454	y4
PIAS3	VSELQVLLGFAGR	694.8984	507.2668	y5
PIAS3	VSELQVLLGFAGR	694.8984	620.351	y6
PIAS3	VSELQVLLGFAGR	694.8984	733.435	y7
PIAS3	VSELQVLLGFAGR	694.8984	832.5034	y8
PIAS3	VSELQVLLGFAGR	694.8984	960.562	y9
PIAS3	VSELQVLLGFAGR	694.8984	1073.646	y10
PIAS3	VSELQVLLGFAGR	694.8984	1202.689	y11
PIAS3	VSELQVLLGFAGR	694.8984	1289.721	y12
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	313.1852	y3
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	460.2537	y4
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	517.2751	y5
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	630.3592	y6
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	743.4432	y7
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	842.5117	y8
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	970.5703	y9
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	1083.654	y10
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	1212.697	y11
PIAS3	VSELQVLLGFAGR[HeavyR]	699.9025	1299.729	y12

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
PIAS4	SYSVALYLVR	585.8295	387.2709	y3
PIAS4	SYSVALYLVR	585.8295	550.3342	y4
PIAS4	SYSVALYLVR	585.8295	663.4183	y5
PIAS4	SYSVALYLVR	585.8295	734.4554	y6
PIAS4	SYSVALYLVR	585.8295	833.5238	y7
PIAS4	SYSVALYLVR	585.8295	920.5558	y8
PIAS4	SYSVALYLVR	585.8295	1083.619	y9
PIAS4	SYSVALYLVR[HeavyR]	590.8336	397.2791	y3
PIAS4	SYSVALYLVR[HeavyR]	590.8336	560.3425	y4
PIAS4	SYSVALYLVR[HeavyR]	590.8336	673.4265	y5
PIAS4	SYSVALYLVR[HeavyR]	590.8336	744.4637	y6
PIAS4	SYSVALYLVR[HeavyR]	590.8336	843.532	y7
PIAS4	SYSVALYLVR[HeavyR]	590.8336	930.5641	y8
PIAS4	SYSVALYLVR[HeavyR]	590.8336	1093.627	y9
PIAS4	LDPDSEIATTGVR	687.3489	331.2083	y3
PIAS4	LDPDSEIATTGVR	687.3489	432.256	y4
PIAS4	LDPDSEIATTGVR	687.3489	533.3036	y5
PIAS4	LDPDSEIATTGVR	687.3489	604.3408	y6
PIAS4	LDPDSEIATTGVR	687.3489	717.4248	y7
PIAS4	LDPDSEIATTGVR	687.3489	846.4674	y8
PIAS4	LDPDSEIATTGVR	687.3489	933.4995	y9
PIAS4	LDPDSEIATTGVR	687.3489	1048.526	y10
PIAS4	LDPDSEIATTGVR	687.3489	1145.579	y11
PIAS4	LDPDSEIATTGVR	687.3489	1260.606	y12
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	341.2166	y3
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	442.2642	y4
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	543.3119	y5
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	614.349	y6
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	727.4331	y7
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	856.4757	y8
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	943.5077	y9
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	1058.535	y10

PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	1155.587	y11
PIAS4	LDPDSEIATTGVR[HeavyR]	692.3531	1270.614	y12

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
PPARγ	HLYDSYIK	519.7663	423.2596	y3
PPARγ	HLYDSYIK	519.7663	510.2917	y4
PPARγ	HLYDSYIK	519.7663	625.3186	y5
PPARγ	HLYDSYIK	519.7663	788.382	y6
PPARγ	HLYDSYIK[HeavyK]	523.7734	431.2738	y3
PPARγ	HLYDSYIK[HeavyK]	523.7734	518.3058	y4
PPARγ	HLYDSYIK[HeavyK]	523.7734	633.3328	y5
PPARγ	HLYDSYIK[HeavyK]	523.7734	796.3961	y6
PPARγ	VEPASPPYYSEK	683.8298	363.1869	y3
PPARγ	VEPASPPYYSEK	683.8298	526.2502	y4
PPARγ	VEPASPPYYSEK	683.8298	689.3135	y5
PPARγ	VEPASPPYYSEK	683.8298	786.3663	y6
PPARγ	VEPASPPYYSEK	683.8298	883.4191	y7
PPARγ	VEPASPPYYSEK	683.8298	970.4511	y8
PPARγ	VEPASPPYYSEK	683.8298	1041.488	y9
PPARγ	VEPASPPYYSEK	683.8298	1138.541	y10
PPARγ	LNHPESSQLFAK	685.8567	365.2178	y3
PPARγ	LNHPESSQLFAK	685.8567	478.3018	y4
PPARγ	LNHPESSQLFAK	685.8567	606.3604	y5
PPARγ	LNHPESSQLFAK	685.8567	693.3925	y6
PPARγ	LNHPESSQLFAK	685.8567	780.4245	y7
PPARγ	LNHPESSQLFAK	685.8567	909.4671	y8
PPARγ	LNHPESSQLFAK	685.8567	1006.52	y9
PPARγ	LNHPESSQLFAK	685.8567	1143.579	y10
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	371.2011	y3
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	534.2644	y4
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	697.3278	y5
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	794.3805	y6
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	891.4333	y7
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	978.4653	y8
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	1049.502	y9
PPARγ	VEPASPPYYSEK[HeavyK]	687.8369	1146.555	y10
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	373.232	y3
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	486.316	y4
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	614.3746	y5
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	701.4067	y6
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	788.4387	y7
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	917.4813	y8
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	1014.534	y9
PPARγ	LNHPESSQLFAK[HeavyK]	689.8638	1151.593	y10
PPARγ	SVEAVQEITEYAK	733.8722	381.2127	y3
PPARγ	SVEAVQEITEYAK	733.8722	510.2553	y4
PPARγ	SVEAVQEITEYAK	733.8722	611.303	y5
PPARγ	SVEAVQEITEYAK	733.8722	724.387	y6
PPARγ	SVEAVQEITEYAK	733.8722	853.4296	y7
PPARγ	SVEAVQEITEYAK	733.8722	981.4882	y8
PPARγ	SVEAVQEITEYAK	733.8722	1080.557	y9
PPARγ	SVEAVQEITEYAK	733.8722	1151.594	y10
PPARγ	SVEAVQEITEYAK	733.8722	1280.636	y11
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	389.2269	y3
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	518.2695	y4
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	619.3171	y5
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	732.4012	y6
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	861.4438	y7
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	989.5024	y8

PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	1088.571	y9
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	1159.608	y10
PPARγ	SVEAVQEITEYAK[HeavyK]	737.8793	1288.651	y11

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
RARG	GLGQPDLPK	462.761	357.2491	y3
RARG	GLGQPDLPK	462.761	472.276	y4
RARG	GLGQPDLPK	462.761	569.3288	y5
RARG	GLGQPDLPK	462.761	697.3874	y6
RARG	GLGQPDLPK	462.761	754.4088	y7
RARG	GLGQPDLPK	462.761	867.4929	y8
RARG	GLGQPDLPK[HeavyK]	466.7681	365.2633	y3
RARG	GLGQPDLPK[HeavyK]	466.7681	480.2902	y4
RARG	GLGQPDLPK[HeavyK]	466.7681	577.343	y5
RARG	GLGQPDLPK[HeavyK]	466.7681	705.4016	y6
RARG	GLGQPDLPK[HeavyK]	466.7681	762.423	y7
RARG	GLGQPDLPK[HeavyK]	466.7681	875.5071	y8
RARG	LQEPLLEALR	591.348	359.2396	y3
RARG	LQEPLLEALR	591.348	488.2822	y4
RARG	LQEPLLEALR	591.348	601.3662	y5
RARG	LQEPLLEALR	591.348	714.4503	y6
RARG	LQEPLLEALR	591.348	811.5031	y7
RARG	LQEPLLEALR	591.348	940.5457	y8
RARG	LQEPLLEALR	591.348	1068.604	y9
RARG	VQLDLGLWDK	593.8269	448.2185	y3
RARG	VQLDLGLWDK	593.8269	561.3026	y4
RARG	VQLDLGLWDK	593.8269	618.324	y5
RARG	VQLDLGLWDK	593.8269	731.4081	y6
RARG	VQLDLGLWDK	593.8269	846.4351	y7
RARG	VQLDLGLWDK	593.8269	959.5191	y8
RARG	VQLDLGLWDK	593.8269	1087.578	y9
RARG	LQEPLLEALR[HeavyR]	596.3521	369.2479	y3
RARG	LQEPLLEALR[HeavyR]	596.3521	498.2904	y4
RARG	LQEPLLEALR[HeavyR]	596.3521	611.3745	y5
RARG	LQEPLLEALR[HeavyR]	596.3521	724.4586	y6
RARG	LQEPLLEALR[HeavyR]	596.3521	821.5114	y7
RARG	LQEPLLEALR[HeavyR]	596.3521	950.554	y8
RARG	LQEPLLEALR[HeavyR]	596.3521	1078.613	y9
RARG	VQLDLGLWDK[HeavyK]	597.834	456.2327	y3
RARG	VQLDLGLWDK[HeavyK]	597.834	569.3168	y4
RARG	VQLDLGLWDK[HeavyK]	597.834	626.3383	y5
RARG	VQLDLGLWDK[HeavyK]	597.834	739.4223	y6
RARG	VQLDLGLWDK[HeavyK]	597.834	854.4492	y7
RARG	VQLDLGLWDK[HeavyK]	597.834	967.5333	y8
RARG	VQLDLGLWDK[HeavyK]	597.834	1095.592	y9

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
RXRα	VLTELVSK	444.7736	333.2127	y3
RXRα	VLTELVSK	444.7736	446.2968	y4
RXRα	VLTELVSK	444.7736	575.3394	y5
RXRα	VLTELVSK	444.7736	676.387	y6
RXRα	VLTELVSK	444.7736	789.4711	y7
RXRα	VLTELVSK[HeavyK]	448.7807	341.2269	y3
RXRα	VLTELVSK[HeavyK]	448.7807	454.311	y4
RXRα	VLTELVSK[HeavyK]	448.7807	583.3536	y5
RXRα	VLTELVSK[HeavyK]	448.7807	684.4012	y6
RXRα	VLTELVSK[HeavyK]	448.7807	797.4853	y7
RXRα	AIVLFNPDSK	552.3084	349.1712	y3

RXRα	AIVLFNPDSK	552.3084	446.224	y4
RXRα	AIVLFNPDSK	552.3084	560.2669	y5
RXRα	AIVLFNPDSK	552.3084	707.3353	y6
RXRα	AIVLFNPDSK	552.3084	820.4194	y7
RXRα	AIVLFNPDSK	552.3084	919.4878	y8
RXRα	AIVLFNPDSK	552.3084	1032.572	y9
RXRα	AIVLFNPDSK[HeavyK]	556.3155	357.1854	y3
RXRα	AIVLFNPDSK[HeavyK]	556.3155	454.2382	y4
RXRα	AIVLFNPDSK[HeavyK]	556.3155	568.2811	y5
RXRα	AIVLFNPDSK[HeavyK]	556.3155	715.3495	y6
RXRα	AIVLFNPDSK[HeavyK]	556.3155	828.4336	y7
RXRα	AIVLFNPDSK[HeavyK]	556.3155	927.502	y8
RXRα	AIVLFNPDSK[HeavyK]	556.3155	1040.586	y9
RXRα	ILEAELAVEPK	606.3477	373.2076	y3
RXRα	ILEAELAVEPK	606.3477	472.276	y4
RXRα	ILEAELAVEPK	606.3477	543.3131	y5
RXRα	ILEAELAVEPK	606.3477	656.3972	y6
RXRα	ILEAELAVEPK	606.3477	785.4398	y7
RXRα	ILEAELAVEPK	606.3477	856.4769	y8
RXRα	ILEAELAVEPK	606.3477	985.5195	y9
RXRα	ILEAELAVEPK	606.3477	1098.604	y10
RXRα	ILEAELAVEPK[HeavyK]	610.3548	381.2218	y3
RXRα	ILEAELAVEPK[HeavyK]	610.3548	480.2902	y4
RXRα	ILEAELAVEPK[HeavyK]	610.3548	551.3273	y5
RXRα	ILEAELAVEPK[HeavyK]	610.3548	664.4114	y6
RXRα	ILEAELAVEPK[HeavyK]	610.3548	793.454	y7
RXRα	ILEAELAVEPK[HeavyK]	610.3548	864.4911	y8
RXRα	ILEAELAVEPK[HeavyK]	610.3548	993.5337	y9
RXRα	ILEAELAVEPK[HeavyK]	610.3548	1106.618	y10
RXRα	GLSNPAEVEALR	628.3356	359.2396	y3
RXRα	GLSNPAEVEALR	628.3356	488.2822	y4
RXRα	GLSNPAEVEALR	628.3356	587.3506	y5
RXRα	GLSNPAEVEALR	628.3356	716.3932	y6
RXRα	GLSNPAEVEALR	628.3356	787.4303	y7
RXRα	GLSNPAEVEALR	628.3356	884.483	y8
RXRα	GLSNPAEVEALR	628.3356	998.526	y9
RXRα	GLSNPAEVEALR	628.3356	1085.558	y10
RXRα	GLSNPAEVEALR	628.3356	1198.642	y11
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	369.2479	y3
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	498.2904	y4
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	597.3589	y5
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	726.4014	y6
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	797.4385	y7
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	894.4913	y8
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	1008.534	y9
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	1095.566	y10
RXRα	GLSNPAEVEALR[HeavyR]	633.3398	1208.65	y11

Transcription Factor	Sequence	Parent ion mass m/z [M+2H]2+	Product ion mass m/z [M+H]+	Transition y-ion
SMAD2	GWGAEYR	419.6957	338.1817	y2
SMAD2	GWGAEYR	419.6957	467.2243	y3
SMAD2	GWGAEYR	419.6957	538.2614	y4
SMAD2	GWGAEYR	419.6957	595.2829	y5
SMAD2	GWGAEYR	419.6957	781.3622	y6
SMAD2	GWGAEYR[HeavyR]	424.6998	348.19	y2
SMAD2	GWGAEYR[HeavyR]	424.6998	477.2326	y3
SMAD2	GWGAEYR[HeavyR]	424.6998	548.2697	y4
SMAD2	GWGAEYR[HeavyR]	424.6998	605.2912	y5
SMAD2	GWGAEYR[HeavyR]	424.6998	791.3705	y6

SMAD2	VETPVLPPVLVPR	708.4346	371.2396	y3
SMAD2	VETPVLPPVLVPR	708.4346	583.392	y5
SMAD2	VETPVLPPVLVPR	708.4346	680.4448	y6
SMAD2	VETPVLPPVLVPR	708.4346	777.4976	y7
SMAD2	VETPVLPPVLVPR	708.4346	890.5817	y8
SMAD2	VETPVLPPVLVPR	708.4346	989.6501	y9
SMAD2	VETPVLPPVLVPR	708.4346	1086.703	y10
SMAD2	VETPVLPPVLVPR	708.4346	1187.75	y11
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	381.2479	y3
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	494.3319	y4
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	593.4003	y5
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	690.4531	y6
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	787.5059	y7
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	900.5899	y8
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	999.6583	y9
SMAD2	VETPVLPPVLVPR[HeavyR]	713.4388	1096.711	y10